

Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 26 2005, 13:45:55 ; Search time 44 Seconds
(without alignments)
900.878 Million cell updates/sec

Title: US-10-721-553-2

Perfect score: 2764

Sequence: 1 MAPTTQTQQRQEDGHPRPNH.....QEDGSEAAAADSSEADSDSD 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgm2_6/_ptodata/1/iaa/5A_COMBO.dep:**
2: /cgm2_6/_ptodata/1/iaa/5B_COMBO.dep:**
3: /cgm2_6/_ptodata/1/iaa/6A_COMBO.dep:**
4: /cgm2_6/_ptodata/1/iaa/6B_COMBO.dep:**
5: /cgm2_6/_ptodata/1/iaa/PCNTUS_COMBO.dep:**
6: /cgm2_6/_ptodata/1/iaa/backfile1.dep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1
US-08-933-750C-9
Sequence 9, Application US/08933750C
; Patent No. 5932442

/022

GENERAL INFORMATION:

APPLICANT: Lal, Preeti Hillman, Jennifer L.
APPLICANT: Baudman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yus, Henry
APPLICANT: Ghegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ FOR WINDOWS Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PITUNOR01

Result No.	Query Score	Query Match Length	DB ID	Description
1	2764	100.0	531	Sequence 9, Appli
2	2764	100.0	531	Sequence 9, Appli
3	2764	100.0	531	Sequence 2, Appli
4	595	21.5	115	Sequence 7407, Ap
5	266	9.6	481	Sequence 18683, A
6	234.5	8.5	1742	Sequence 4, Appli
7	234.5	8.5	1742	Sequence 4, Appli
8	233.5	8.4	930	Sequence 3, Appli
9	226	8.2	918	Sequence 1, Appli
10	225	8.2	933	Sequence 2, Appli
11	225.5	8.2	933	Sequence 2, Appli
12	225.5	8.2	936	Sequence 5249, Ap
13	225.5	8.2	936	Sequence 5249, Ap
14	223.5	8.1	1315	Sequence 5, Appli
15	220.5	8.0	1259	Sequence 10366, A
16	218	7.9	1166	Sequence 7, Appli
17	212.5	7.7	287	Sequence 468, App
18	212.5	7.7	1092	Sequence 15, Appli
19	203.5	7.4	414	Sequence 19046, A
20	203	7.3	257	Sequence 188, App
21	203	7.3	272	Sequence 186, App
22	199.5	7.2	599	Sequence 864, App
23	198.5	7.2	238	Sequence 190, App
24	196	7.1	781	Sequence 9773, Ap
25	195	7.1	231	Sequence 194, App
26	195	7.1	232	Sequence 192, App
27	195	7.1	764	Sequence 944, App

Query Match	1 MAPTICQ
Best Local Similar Matches	1 MAPTIQ
Matches 53+; Conn.	61 YKATSSL
	61 YKATSSL
	121 SQHQAKY
	121 SQHQAKY
Qy	241 MSQAMII
Db	241 MSQAMII
Qy	301 NKASKGII
Db	301 NKASKGII
Qy	361 LEQDEAI
Db	361 LEQDEAI
Qy	421 EEGDRDI
Db	421 EEGDRDI
Qy	481 DSDBGSII
Db	481 DSDBGSII

RESULT 2
US-09-234-613-9
Sequence 9, Application
Patient No. 6132373
GENERAL INFORMATION
APPLICANT: Lal Hill
APPLICANT: Bandy
APPLICANT: Shah Al
APPLICANT: Yue
APPLICANT: Gue
APPLICANT: Cor
TITLE OF INVENTION
NUMBER OF SEQUEL
CORRESPONDENCE
ADDRESSEE: I
STREET: 3174
CITY: Palo A
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE
MEDIUM TYPE:
COMPUTER: IBM
OPERATING SYSTEM:
SOFTWARE: Fair
CURRENT APPLICANT
APPLICATION N

FILING DATE:				
CLASSIFICATION:				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER:	US/08/933,750			
FILING DATE:	September 23, 1997			
ATTORNEY/AGENT INFORMATION:				
NAME:	BILLINGS, LUCY J.			
REGISTRATION NUMBER:	36,749			
REFERENCE/DOCKET NUMBER:	PF-0356 US			
TELECOMMUNICATION INFORMATION:				
TELEPHONE:	415-855-0555			
TELEFAX:	415-845-4166			
TELEX:				
SEQUENCE FOR SEQ ID NO: 9:				
SEQUENCE CHARACTERISTICS:				
LENGTH:	531 amino acids			
TYPE:	amino acid			
STRANDEDNESS:	single			
TOPOLOGY:	linear			
IMMEDIATE SOURCE:				
LIBRARY:	PTUNOR01			
CLONE:	98974			
US-09-234-613-9				
Query Match	100.0%	Score: 2754;	DB: 3;	Length: 531;
Best Local Similarity	100.0%	Pred. No. 4.2e-227;		
Matches 531;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	M A P T I O T Q A Q R E D G H R P N S H R T L P P R S G Y C R V K C N S L P D I P P P K F I T Y P F D Q N R F V Q	60	
Db	1	M A P T I O T Q A Q R E D G H R P N S H R T L P P R S G Y C R V K C N S L P D I P P P K F I T Y P F D Q N R F V Q	60	
Qy	61	Y K A T S L E K O H K H D L L T E P D I G V T I D I N P D P T Y R I D P N V I L D P A D E K L L E B I Q A T S S K R	120	
Db	61	Y K A T S L E K O H K H D L L T E P D I G V T I D I N P D P T Y R I D P N V I L D P A D E K L L E B I Q A T S S K R	120	
Qy	121	S Q Q H A K V P W M R K T K E T I S T E F N R Y C T S N E P K V I G V S T K Q Q F T E E I Y K D R D S C T I A T E	180	
Db	121	S Q Q H A K V P W M R K T K E T I S T E F N R Y C T S N E P K V I G V S T K Q Q F T E E I Y K D R D S C T I A T E	180	
Qy	181	K T P F D Q K S I S Q H Y S K P R T P V E V M P V P D F K M N I N P C A Q V I F D S D P A P K D T S G A A L E M	240	
Db	181	K T P F D Q K S I S Q H Y S K P R T P V E V M P V P D F K M N I N P C A Q V I F D S D P A P K D T S G A A L E M	240	
Qy	241	M S Q A M I R G M M D B E G N O F V A Y F L P V E T I L K G K R D B E E M D Y A P D D V D Y K I K A R E R T N W V K	300	
Db	241	M S Q A M I R G M M D B E G N O F V A Y F L P V E T I L K G K R D B E E M D Y A P D D V D Y K I K A R E R T N W V K	300	
Qy	301	N K A S K G Y E E Y N F F I I P R E G D G V Y Y N E L E T R Y L S K R A K A V Q S G T N A L L V V K H R M N N E K E	360	
Db	301	N K A S K G Y E E Y N F F I I P R E G D G V Y Y N E L E T R Y L S K R A K A V Q S G T N A L L V V K H R M N N E K E	360	
Qy	361	L E A Q E A R K A Q L E N H I R E E E E M E T E E K A G G S D E E Q K G S S S K E G S E D E H S G S E S E R	420	
Db	361	L E A Q E A R K A Q L E N H I R E E E E M E T E E K A G G S D E E Q K G S S S K E G S E D E H S G S E S E R	420	
Qy	421	E E G D R D E A S D K S G S G D E S S E D E A R A A R D K E I F P S D A D S D D E R G Q A Q G S D N	480	
Db	421	E E G D R D E A S D K S G S G D E S S E D E A R A A R D K E I F P S D A D S D D E R G Q A Q G S D N	480	
Qy	481	D S D S G N G G G O R S R S H S R S A P S P G S E S H A A Q D G S E A A D S D S A D S D S D	531.	
Db	481	D S D S G N G G G O R S R S H S R S A P S P G S E S H A A Q D G S E A A D S D S A D S D S D	531.	

Page 3

TITLE OF INVENTION: No. 6680196 gel Gene That is Amplified and
 Title of Invention: Overexpressed in Cancer and Methods of Use Thereof
 FILE REFERENCE: UNMC6121
 CURRENT APPLICATION NUMBER: US/09/647,143
 CURRENT FILING DATE: 2000-09-27
 PRIORITY NUMBER: PCT/US99/066533
 PRIOR FILING DATE: 1999-03-26
 PRIOR APPLICATION NUMBER: 60/079,649
 PRIOR FILING DATE: 1998-03-27
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 2
 LENGTH: 531
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-647-143-2

Query Match		Score 2764; DB 4; Length 531;	
Best Local Similarity		Pred. No. 4.2e-227; Indels 0; Gaps 0	
Matches 531; Conservative		Mismatches 0; Insertions 0; Deletions 0	
Qy	1 MAPIOTQQAQEDGHRENSHTLPERSGVYCRVKCNSLPDIPEFDPKITYPFQDNQRFVQ	Db	1 MAPIOTQQAQEDGHRENSHTLPERSGVYCRVKCNSLPDIPEFDPKITYPFQDNQRFVQ
Qy	61 YKATSLLEKQHKHDLIIPEDLGYTIDLNPDTRIDPNVLLPADEKLEERIQAQPTSSKR	Db	61 YKATSLLEKQHKHDLIIPEDLGYTIDLNPDTRIDPNVLLPADEKLEERIQAQPTSSKR
Qy	121 SQQHAKVVPWNRKTEVISTEFNRYGINSNEKEPVKGIVSYVKQOFTEEIIYKORDSQITAIE	Db	121 SQQHAKVVPWNRKTEVISTEFNRYGINSNEKEPVKGIVSYVKQOFTEEIIYKORDSQITAIE
Qy	181 KTFPEDAQKSISQHYSIQRPTVPEVMPVFPDPFQWINCAQVIFDSDAPKDTSQAALEM	Db	181 KTFPEDAQKSISQHYSIQRPTVPEVMPVFPDPFQWINCAQVIFDSDAPKDTSQAALEM
Qy	241 MSQAMIRGMMDDEGNDFVAYFLPVESTLKKRKDQEEMDYAPDDVYDKAREYWNVK	Db	241 MSQAMIRGMMDDEGNDFVAYFLPVESTLKKRKDQEEMDYAPDDVYDKAREYWNVK
Qy	301 NKASKCYEENYFFIEGDGTYYNELETRVLSKRRRAKAGVQSGTNALLYVXKHDRDNKEK	Db	301 NKASKCYEENYFFIEGDGTYYNELETRVLSKRRRAKAGVQSGTNALLYVXKHDRDNKEK
Qy	361 LEAQEARKAQLENHSPBEEEEMETEEKEAGGSDEEQEKGSSEKEGSSEDEHSGSESER	Db	361 LEAQEARKAQLENHSPBEEEEMETEEKEAGGSDEEQEKGSSEKEGSSEDEHSGSESER
Qy	421 EGGDRDEASDKSGSFDESSDEARARDKEIFGSDADSDDEDDQAOQGSDN	Db	421 EGGDRDEASDKSGSFDESSDEARARDKEIFGSDADSDDEDDQAOQGSDN
Qy	481 DSDSGSNGGGRSRHSRSASPPGSEHSAQEDGSEAAASDSSEADSDSD 531	Ddb	481 DSDSGSNGGGRSRHSRSASPPGSEHSAQEDGSEAAASDSSEADSDSD 531

RESULT 4
 US-09-513-999C-7407
 Sequence 7407 Application US/09/13999C
 Patient No. 6783961
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, J.B.
 APPLICANT: Duclert, A.
 APPLICANT: Giordano, J.Y.
 TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 Patient No. 6783961
 FILE REFERENCE: 59-US2-REG
 CURRENT APPLICATION NUMBER: US/09/513,999C
 CURRENT FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/122,487
 PRIOR FILING DATE: 1999-02-26

```

NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent-pm
SEQ ID NO: 7407
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens

; FEATURE;
; NAME/KEY: UNSURB
; LOCATION: 25
; OTHER INFORMATION: Xaa=Glu
; FEATURE;
; NAME/KEY: UNSURE
; LOCATION: 26
; OTHER INFORMATION: Xaa=Arg
; FEATURE;
; NAME/KEY: UNSURE
; LOCATION: 11.0
; OTHER INFORMATION: Xaa=Glu
; FEATURE;
; NAME/KEY: UNSURE
; LOCATION: 11.4
; OTHER INFORMATION: Xaa=Ala
US-09-513-399C-7407

Query Match 21.5
Best Local Similarity 96.5
Matches 111; Conservative

QY          1 M A P T I Q T O A Q R E D G H R R
Db          1 M A P T I Q T O A Q R E D G H R R
QY          61 Y K A T S L E K Q H K H D L L T
Db          61 Y K A T S L E K Q H K H D L L T

RESULT 5
US-09-248-796A-18683
; Sequence 18683, Application US
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al.
; TITLE OF INVENTION: NUCLEIC
; TITLE OF INVENTION: FOR DIA-
; FILE REFERENCE: 107196_132
; CURRENT APPLICATION NUMBER:
; CURRENT FILING DATE: 1999-01-01
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO: 18683
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18683

```

136 EPB ----- vsflrrtevis ----- DRPLSKGGNNLNTATEEIKVKE 171
 QFTEEEIYKORDSQTIAEKTFDAQKST-----SQHYSKPRVTPTVMPVFPDFKWIN 216
 KLSDEHF-DADSDQNTVANESFYKHNKPKKKHLRAVNTWPLJPDTSMLDN 230
 PCAQVF-DSDPDKTSSAALEMSOMIRNMDE-----EGNQFV 258
 VPINLRFGMASINRELNNKLQQQQQQQNDKPKFLEGGEWI 290
 AYFL------PVEETLCKRKDQEEMDYAPDDYD-YKTAERYNNWV-----kNKA 303
 SMYLDATNTSTTANDNDEQ1-----NLVEKLHSTKKEQPINLDBEESLETVKFKY 346
 SKGBENYFPIFPEGDGV-----YTNELTRVLSKRRA- 337
 TKNYDMTCOPFHENEELAIKFVSDEIEPVSKDNFKRKRNAYYTPINGKIELKGRAS 406
 KAGVSGTNALLVVKHRDNMKELEAQARAKAOLNHEPEEEEMETE 388
 TNSPINKPTKERTDGTMFL---RESTTNBLKRLDTIRSEYDPMHEYGEDEBEEEEE 462
 KEAGGSDEEQEGSSSRK 407
 EEEPLEEEQQQETSTKEE 481

RESULT 6
 US-09-386-962C-4
 Sequence 4, Application US/09386962C
 Patent No. 6635473
 GENERAL INFORMATION:
 APPLICANT: FOSTER, Timothy J.
 TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLO
 FILE REFERENCE: P06335USP/BAS
 CURRENT APPLICATION NUMBER: US/09/386,962C
 CURRENT FILING DATE: 199-08-31
 PRIOR APPLICATION NUMBER: 60/098,443
 PRIOR FILING DATE: 1998-08-31
 PRIOR APPLICATION NUMBER: 60/117,119
 PRIOR FILING DATE: 1999-01-25
 NUMBER OF SEQ ID NOS: 38
 SEQ ID NO 4
 LENGTH: 1742
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis

RESULT 7
 US-09-386-959-4
 Sequence 4, Application US/09386959
 Patent No. 6703025
 GENERAL INFORMATION:
 APPLICANT: PATTI, Joseph M.
 APPLICANT: FOSTER, Timothy J.
 APPLICANT: HOOK, Magnus
 TITLE OF INVENTION: MULTICOMPONENT VACCINES
 FILE REFERENCE: P06333US1/BAS
 CURRENT APPLICATION NUMBER: US/09/386,959
 CURRENT FILING DATE: 1999-08-31
 EARLIER APPLICATION NUMBER: 60/098,439
 EARLIER FILING DATE: 1999-08-31
 NUMBER OF SEQ ID NOS: 65
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 4
 LENGTH: 1742
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis

Query Match Score 234.5; DB 4; Length 1742;
 Best Local Similarity 21.0%; Pred. No. 8.5e-11; Mismatches 88; Indels 149; Gaps 24;
 Matches 127; Conservative 88; MisMatches 240; InDelS 149; Gaps 24;

48 FITYPDQNRQFVQYKATSLEKOKHDLITPDLGVTIDLNPDYRIDL-----NVLLDP 102
 YVTLKDSNRELQRTTDQSGHYQFDNLQNGT-YTVEFATEDDNYTPSPANNSNDAIDS 773
 103 ADEKLLEEQIQTSSKRSQHAKV-----VWMRRTKEYISTEFNRYGISENEKEVKIG 156
 DGERDGTRKVVAKGTINNADMTVDTGFLTPKVNQD-----KRDQS 175
 VSV-----KOQFT-EETTY-----KRDQS 175
 VKYTLKNGQDITSTTDSNGYKEPTSIENQTYTIEFTPEGYTPKQNSGSDEKDSDN 893
 ITAEKTFEDA-OKSIOSHKSPRVTPTVMPVFPDFKWINCAQVIFDSDP-----227
 GKTPTVKAQGTTDQSGHYQFDNLQNGT-----LGDI-WEDDTNKDGCIQDSBEKG1SGV 945
 APKDTSGAA-ALEMSQAMIRGMMDDEGQFVAYFLPVEETLKKRKDQEEEMDY- 281
 VTLKDKNGNAIGTTDASGHYQFKGL-ENGSYTVEFPTSGYTPTRANSGODITVDSN 1003
 282 -----APDDYD-YKIA-----BYNWVYRNK-----ASGY-----307

1004 GITTGTINGADNTLDSGFVTPKSVGVDYWDNTNKOGIQQDNBKGJISGVVKTLDEK 1063
 Db Qy 281 -----YAP-----DDVYD-----
 Qy 546 YQFTGLNSGTSYVBSTPAGYPTTANGTDDAVSCLLTGVIKADNMFLDSGFVYKT 605
 Db Db 292 AR --- EYNWYTKNK --- -ASKGY-----
 Qy 605 PKYSLGDYIWYDENKGKRDSTEGIKVVKTLQNEKGEVIGTETDNGKTRFDNLDSG 665
 Db 358 EKELEAQEARKA ----- OLENHPEPEEREEEMTEKEAGGSDEEQEKGSSSEKEGSBE 410
 Db 1110 ERDAGEDEVRVITIDDDFSIDNGYEDDDSDSDSDSDSDSDSDSDSDSDSDSDA 1169
 Qy 411 DEHGSBESERE-FGDRDEASDK-SGSGEDDSSDEARAARDKEEIFGSDADSDDDADSD 468
 Db 1170 DSD 1229
 Qy 469 EDRGQAQGGSNDSGSNGGGQRSRSHSSASPPPGSHEHSQAEDGSAADS-SEAD 527
 Db 1230 SDSDDSD 1286
 Qy 528 SDS 531
 Db 1287 SDS 1290

RESULT 8
 US-09-200-650E-3
 ; Sequence 3, Application US/09200650E
 ; Patent No. 6680195
 ; GENERAL INFORMATION:
 ; APPLICANT: Patti, Joseph M.
 ; APPLICANT: Foster, Timothy J.
 ; APPLICANT: Hook, Magnus A.O.
 ; APPLICANT: Deihinn, Deirdre Ni.
 ; APPLICANT: Perkins, Samuel L.
 ; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
 ; FILE REFERENCE: P06231US2/BAS
 ; CURRENT APPLICATION NUMBER: US/09/200, 650E
 ; CURRENT FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: 60/006, 815
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR FILING DATE: 1998-08-31
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 3
 ; LENGTH: 930
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; US-09-200-650E-3

Query Match Score 233 5; DB 4; Length 930;
 Best Local Similarity 20 4%; Pred. No. 4.2e-11; Gaps 25;
 Matches 130; Conservative 20 5%; Mismatches 211; Indels 215; Gaps 25;

Qy 53 FDQNRFVQYKATSIDEKHADLLSPDLGTTIDLNPDYTRIDPNVLLDADEKLLEB 112
 Db 298 FEQVAFKRKNATDK-----TAKMETY--LGNDTY-----SBEI 331
 Qy 113 QAPTSKRSOQHNTVPMMEKTYSTE-----
 Db 332 IVYGNKKAQ -----PLISINTINNEDLSRNMATTAYNQPKNTYTKTFVTNLTYKFN 385
 Qy 141 -----FNRYGTISNEK-----PEVKIGVSKVQOFTEEYKQDTSQITAIEKFEDAQK 188
 Db 386 PNAKNPFIYVTDQFVDFSTPDTSKLKQVTDQF---DVTYSNDNKTATVD-LMKQQT 440
 Qy 189 SISQHYSKPKRTPVHEMPVPDFKOMWINPAQVIFPDSDPAPKDTGAALEMMSQAMIRG 248
 Db 441 SSNKQYIILQGY-----AVPDNSSTTDNGKIDYTLDTKTKYWSN-----SYSNYNG 486
 Qy 249 MMDBEBNQFAYFL --PVSETLKKRKDDE-----EMD-----280
 Db 487 SSTANGDQ-KKYNLLEDYWWEDNTKDGKODANEKGIKGVTVILKDSNGKELDRTTDENGK 545

Query Match Score 226; DB 4; Length 918;
 Best Local Similarity 22 6%; Pred. No. 1.8e-10;
 Matches 125; Conservative 75; Mismatches 208; Indels 146; Gaps 23;

Qy 33 VRYCNDSLDPIPP-DPKITYPFDQNRVQYKATSLEXKOHKDILTEPDGVYIDLIN-- 88
 Db 276 VDYSNSNNTMP1ADIK-----STNGDVVARAT-----IDLTKTYTFTVTDVNRK 322
 Qy 89 -----PDTYRIDPNVLLDADEKLLE-----EIOADPTSSKRS 121
 Db 323 NINGQFSLPLFTDRAKPKSGTYDANI-----ADEMFNFKITYSSPIAGIDKPGNANTS 380
 Qy 122 QQHAKTVWPARTB1IATEFNRGIGNSKPKFVIGSVKQOFTEEYKDRSQITAE- 180
 Db 381 SQ1QVNTDASGONTYKQTVF----VNPKVRLGNTWVYIKGQDKI-EESSSKVSTDT 434
 Qy 181 --KTFE--DAQKSISQHYSKPKRVTPV-EVMPVPDFKOMWINPAQVLFDSDPAPKDTSGA 235
 Db 435 KURIFENDTSKLSDSYIADPDSNLKEVTDOKNRYYEHNVAVSKFGD-----485
 Qy 236 AALEMMSOAMIRGMMDDEGNGQFWAYFLPVEETLKKRKRDQEEMDVAPODDVYDYKLAIREY 295

Page 6

Sequence 5249, Application US/08956171E
 Patent No. 6591114
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 Gil H. Choi
 Patrick S. Dillon
 Craig A. Rosen
 Steven C. Barash
 Michael R. Fannon
 TITLE OF INVENTION: *Staphylococcus aureus* Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5256
 CORRESPONDENCE ADDRESS:
 ADDRESS: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MS DOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIORITY DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 5249:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 936 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 5249:

Sequence 5249, Application US/08956171E
 Patent No. 6591114
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 Gil H. Choi
 Patrick S. Dillon
 Craig A. Rosen
 Steven C. Barash
 Michael R. Fannon
 TITLE OF INVENTION: *Staphylococcus aureus* Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5256
 CORRESPONDENCE ADDRESS:
 ADDRESS: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MS DOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIORITY DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 5249:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 936 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 5249:

Query Match 8.2%; Score 225.5; DB 4; Length 936;
 Best Local Similarity 21.4%; Pred. No. 2e-10;
 Matches 103; Conservative 76; Mismatches 187; Indels 115; Gaps 18;

83 TIDLINP--DTTRIDPNVLDPADKEKLEIQAQPTSKRSQOHAKVPPWRKTEYISTE 140
 392 TIDQIDKTNNTR--OTIVYNPSGDNVYI----APVLT-----
 141 FNRYGISNEKPEVYIGVSKVQOETEEYIK----DRDSQTIAKEFEDAKSISDHSY 195
 423 -----GRLKENTDSNALIDQDQNTSIVKYKVDNAADLSSESYFNPNPENFDVTNSVNTP 476
 196 KPRVTPVEMMPVFDFKWNPCAQVIFSDPAPKDTSGAALEM--SQAMIRGM- 249
 477 NPNQYKVEFT- EDDQITTPYTVVNCIDP--NSKGDLALRSTLYGYNNSNIWRSMS 531
 250 MDEEGNQFAYFL----PVEETLKKRKRDQEEMDYAPDDYYDKIARETNWNYKKA 303
 532 WDNB----VAFNNGSGSGDGIKPKVVPQEOPDEGEIPEPD----- SDS 572
 304 SKGYEEVNFPIFRGDGVYNELETRVLSSKERAAGVQSGTNMALL--- VVK 352
 573 DSDNSDSSGGSGDSGDSSTDGSASDSASDS 610

RESULT 13
 US-08-781-986A-5249
 Sequence 5249, Application US/08781986A
 Patent No. 6737248
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 TITLE OF INVENTION: *Staphylococcus aureus* Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5255
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MS DOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,986A
 FILING DATE:
 CLASSIFICATION: 4.35
 PRIORITY DATA:
 PRIORITY NUMBER: PB248P1
 ATTORNEY/AGENT INFORMATION:
 NAME: Benson, Bob
 REGISTRATION NUMBER: 30,446
 REFERENCE/DOCKET NUMBER: PB248PP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEXFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 5249:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 936 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-781-986A-5249
 Query Match 8.2%; Score 225.5; DB 4; Length 936;
 Best Local Similarity 21.4%; Pred. No. 2e-10;
 Matches 103; Conservative 76; Mismatches 187; Indels 115; Gaps 18;
 83 TIDLNP--DTTRIDPNVLDPADKEKLEIQAQPTSKRSQOHAKVPPWRKTEYISTE 140
 392 TIDQIDKTNNTR--OTIVYNPSGDNVYI----APVLT-----
 141 FNRYGISNEKPEVYIGVSKVQOETEEYIK----DRDSQTIAKEFEDAKSISDHSY 195
 423 -----GRLKENTDSNALIDQDQNTSIVKYKVDNAADLSSESYFNPNPENFDVTNSVNTP 476
 196 KPRVTPVEMMPVFDFKWNPCAQVIFSDPAPKDTSGAALEM--SQAMIRGM- 249
 477 NPNQYKVEFT- EDDQITTPYTVVNCIDP--NSKGDLALRSTLYGYNNSNIWRSMS 531
 250 MDEEGNQFAYFL----PVEETLKKRKRDQEEMDYAPDDYYDKIARETNWNYKKA 303
 532 WDNB----VAFNNGSGSGDGIKPKVVPQEOPDEGEIPEPD----- SDS 572
 304 SKGYEEVNFPIFRGDGVYNELETRVLSSKERAAGVQSGTNMALL--- VVK 352
 573 DSDNSDSSGGSGDSGDSSTDGSASDSASDS 610

Qy 196 KPRVTPVTEMPVFDPKRMINPCQVIFSPDPAKDTGAALEM---SQAMIRGM- 249
 Db 477 NPNQKVPEVFNT- PDDQLTTPYIVVNHGIDP--NSKGDALARSTLVGYNSTNIWRSMS 531
 Qy 250 MDEBGNQFVAYPL----PVEETLKKRKDQEEMDAPPDVYDYKLALEYNNWNVKKA 303
 Db 532 WDNE---VAFNNSSGSGDQIKDKEVVPEDPGEIEPTPED-----SDS 572
 Qy 304 SKGYEENYFFIIFRGDGYYNELETRVLISKRAKAGYQSGTNALL----VVK 352
 Db 573 DPGDSG-----SDNSDSDGSQDSDSSDSASASDSDSAS 610
 Qy 353 HRDMNEKELEAQKAKOQNHEDEEEEMETEEKEAGGSDEQEGKSSSEKEGSEDE 412
 Db 611 DSASASDPSASDSDASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 670
 Qy 413 HSSESESE-BGDDDEASDPSGSNEDDESEDEAARDKEEIFSSDADSDSDSDSD 471
 Db 671 DSDDSD 729
 Qy 472 GOAQGGSDNDSDSSNSNGGQRQRSHRSRHSASAPFGSEHSAQEDGSEAARSDS-SEADS 530
 Db 730 SD 788
 Qy 531 D 531
 Db 789 D 789

RESULT 14
 US-09-200-650E-5
 ; Sequence 5 , Application US/09200650E
 ; Patent No. 6680195
 ; GENERAL INFORMATION:
 ; APPLICANT: Patti, Joseph M.
 ; APPLICANT: Foster, Timothy J.
 ; APPLICANT: Hock, Magnus H.O.
 ; APPLICANT: Bidhman, Deirdre Ni.
 ; APPLICANT: Perkins, Samuel L.
 ; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from *Staphylococcus aureus*
 ; CURRENT APPLICATION NUMBER: US/09/200,650E
 ; CURRENT FILING DATE: 1998-11-25
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: 60/098,427
 ; PRIOR FILING DATE: 1998-08-31
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.0
 ; LENGTH: 1315
 ; TYPE: PRT
 ; ORGANISM: *Staphylococcus aureus*
 ; SEQ ID NO: 5
 US-09-200-650E-5

Query Match Score 223.5; DB 4; Length 1315;
 Best Local Similarity 21.7%; Pred. No. 4.9e-10;
 Matches 121; Conservative 85; Mismatches 225; Indels 127; Gaps 23;

Qy 27 SGIVTCRVYCNISLDIPDPKITYPPDORFRYQKATSLKEKHDLITEPDIGTIDL 86
 Db 772 TGTY------NGADNMTLDSGF--YKTPXNLGRNYWWDPTNKDKQDSTEKGSGVTWL 823
 Qy 87 INPD-----TYRID-----PVNLDPADEKLLBEIOP 115
 Db 824 RVENGEULQTTKDQGKXQFTGLENGYKVERPPTSGYPTQGSCDDEG-1DSNGST 882
 Qy 116 TSSKRKRAKAV----VWMRKTKEYISTEFENYGISNEKEPEYKIGSYKQQTTEEETYK 170
 Db 883 TGYIKDONDIDBGFYKPTYLNQDWWDTNKGVQDDEKGIGSVTV-----TLK 934
 Db 171 DRDSQITAIKTFDAAKSISQ----HYSKSPRITPVEIMPFDFKMWINPACQVIFSD 226
 Qy

935 DENDKVLKTTTDDENGKYQFTDLNNGTYKVBETPQGTYPT-----SVTSGN 981
 Db 979 PAPKDTGAALEMMSQAMIRGMMDDEG-----NOFYAVFLPVEETIKKRKDOEE 277
 Qy 982 DTEKDONGLTTGVIKDA- DNMTLUSGFYKTPKSLGQDLYWY-----DSNKDGQDOSTE 1034
 Db 278 EMYAPDDVYDYKLALEYNNWNVKKA- ASKGYEENYFFFIREGDGYYNELETRVLISK 335
 Qy 1035 K----GIXDKVYTL---LINEKGEVIGTKTDENGSKYCFONLDSGXY----KVIFEK- 1079
 Db 336 RAKAGV-QSGTNALIVKRDMMNEKELEAQKAKOQNHEPEEEBEMETEEKAGGS 394
 Qy 1080 -PAGITQGTNTTDEDKDAQGGEVDTVITHDFFLDNGYTYEEETSD-----§ 1126
 Db 395 DEEQEKGSSSEKEGSSEDEHSGSESEREGDRDEASDPSGSDEDESEDEAARDKEBIF 454
 Db 1127 DSDDSD 1184
 Qy 455 GSDADSEDDADSDSDSDERDQGQCGSDNDSDSGNSNGGQRSRSHRSRSPFPSEHSAQED 514
 Db 1185 DSDDSD 1237
 Qy 515 GSEAAADS-SEADS 531
 Db 1238 DSDDSDSDSDSDSDSDSDSD 1255

RESULT 15
 US-09-949-016-10366
 ; Sequence 10366 , Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 10366
 ; LENGTH: 1259
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-10366

Query Match Score 220.5; DB 4; Length 1259;
 Best Local Similarity 23.1%; Pred. No. 8.3e-10;
 Matches 93; Conservative 37; Mismatches 160; Indels 93; Gaps 16;

Qy 171 DRDSQITALEKTPEDAQKSISQHYSKSPRTPVTPVMEVFPDFKMWINPACQVIFDSDPARK 230
 Db 297 DHDSSTI-----GQNSDSKEYDPEGB-----DPHNEV-DGDKTSK 331
 Qy 231 DTSGZAALEMMSQAMIRGMDEEGNQFVAVPLPVESTLK----KR----KRDQEE 278
 Db 332 SEENSA-----GIPDNGSQ-----RIETOQKLNRRESRVNTLTKESETHA 374
 Qy 279 MDYAPDDVYDYKIALEYNNWNVKKA-SKGYE----BNYFFT FREGGVYNE---LETTRV 330
 Db 375 VGKSDQKGIBKGPSGGRNNTIKEVGKNGNEKELEAQKAKOLENHEPEEEEP-----E 382
 Qy 331 RLSKRRAKAVQSGTNALLYVKHRDMNEKELEAQKAKOLENHEPEEEEP-----E 382
 Db 435 QKSEPGKVKG-HSNTGS-----DSNSDGYDSDDNSDSSDINGNDANS 486
 Qy 383 EMETEKEAG----GSDDEOEKGSSSKEESEDNSGSESEREGIR-----DEASDK 431

Db	487	EDSNNSSSRGDASYNDESKONGGSDSGAEDDSSTDNSTNSDSNGNGNNDNDK	546
Qy	432	SGSGE -- DESSDEARARDREEIFGSDADSEDDADSDEDRGQAQGSNDSDSGSNG	488
Db	547	SDSGKGSKSDSSDSDSSDSSNNSSSSDSDSSDSDSSDSDSSDSDSSDSDS-SDS	605
Qy	489	GGQRSRSHRSRASPPGSEHSAQEDGSEAA-SDS-SEA-ADSD	531
Db	606	SNSSDSDSSDSDSSDSDSSKSDSSKSDSSKSDSSKSD	648

Search completed: September 26, 2005, 13:56:26
Job time : 47 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2005, 13:48:56 ; Search time 174 seconds

(without alignments)
 1242.129 Million cell updates/sec

Scoring table: BLOSUM62 Gapext 0.5

Searched: 1826554 seqs, 407025358 residues

Total number of hits satisfying chosen parameters: 1826554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing First 45 summaries

Database : Published Applications AA:^{*}

1: /cgn2_6/_prododata/2/_pubpa/US07_PUBCOMB.pep:*

2: /cgn2_6/_prododata/2/_pubpa/_PCT_NEW_PUB_pep:*

3: /cgn2_6/_prododata/2/_pubpa/_US05_NEW_PUB_pep:*

4: /cgn2_6/_prododata/2/_pubpa/_US06_PUBCOMB.pep:*

5: /cgn2_6/_prododata/2/_pubpa/_US07_NEW_PUB_pep:*

6: /cgn2_6/_prododata/2/_pubpa/_PCTO5_PUBCOMB.pep:*

7: /cgn2_6/_prododata/2/_pubpa/_US08_NEW_PUB_pep:*

8: /cgn2_6/_prododata/2/_pubpa/_US08_PUBCOMB.pep:*

9: /cgn2_6/_prododata/2/_pubpa/_US09_PUBCOMB.pep:*

10: /cgn2_6/_prododata/2/_pubpa/_US09B_PUBCOMB.pep:*

11: /cgn2_6/_prododata/2/_pubpa/_US09C_PUBCOMB.pep:*

12: /cgn2_6/_prododata/2/_pubpa/_US09_NEW_PUB_pep:*

13: /cgn2_6/_prododata/2/_pubpa/_US10A_PUBCOMB.pep:*

14: /cgn2_6/_prododata/2/_pubpa/_US10B_PUBCOMB.pep:*

15: /cgn2_6/_prododata/2/_pubpa/_US10C_PUBCOMB.pep:*

16: /cgn2_6/_prododata/2/_pubpa/_US10D_PUBCOMB.pep:*

17: /cgn2_6/_prododata/2/_pubpa/_US10E_PUBCOMB.pep:*

18: /cgn2_6/_prododata/2/_pubpa/_US10F_PUBCOMB.pep:*

19: /cgn2_6/_prododata/2/_pubpa/_US11A_PUBCOMB.pep:*

20: /cgn2_6/_prododata/2/_pubpa/_US11A_NEW_PUB_pep:*

21: /cgn2_6/_prododata/2/_pubpa/_US10_G_PUBCOMB.pep:*

22: /cgn2_6/_prododata/2/_pubpa/_US10_G_PUBCOMB.pep:*

RESULT 1

US-09-840-787-9

; Sequence 9, Application US/09840787

; Patent No. US2002005264A1

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti; Hillman, Jennifer L.

; Bandman, Olga

; Shah, Purvi

; Au-Young, Janice

; Yue, Henry

; Guegler, Karl J.

; Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/B40,787

FILING DATE: 23-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/518, 865

FILING DATE: <Unknown>

SUMMARIES

Result No.	score	Query Match Length	DB ID	Description
1	2764	100.0	531 9	Sequence 9, Appli
2	2764	100.0	531 17	Sequence 2, Appli
3	2658.5	96.2	553 18	Sequence 50041, A
4	2464	89.1	473 10	Sequence 410, App
5	2464	89.1	473 17	Sequence 410, App
6	622	22.5	133 18	Sequence 50040, A
7	452	16.4	86 15	Sequence 223174, A
8	377.5	13.6	571 15	Sequence 21359,
9	337.5	12.2	644 16	Sequence 116147,
10	283	10.2	475 18	Sequence 49771, A
11	238.5	8.6	286 16	Sequence 202470, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PP-0356 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEX: 415-845-4166
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 531 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: PITTUNDR01
 CLONE: 98974
 SEQUENCE DESCRIPTION: SEQ ID NO: 9 :

US-09-840-787-9

	Query Match	Score	DB 9;	Length
Qy	MAPTIQTQREDFHRPNSHRTLPERSGVTCRKYCNSLPDIIFDPKPIYPPDQI	100.0%	2764;	531;
Db	MAPTIQTQREDFHRPNSHRTLPERSGVTCRKYCNSLPDIIFDPKPIYPPDQI	100.0%	Prod. No. 9.8e-167;	0
Qy	YKATSLEROKHHDLTTEPDGLVTIDLNPDTYRIDPNVLLPADEKLLEEEQAP	61	Mismatches	0; Indels 0;
Db	YKATSLEROKHHDLTTEPDGLVTIDLNPDTYRIDPNVLLPADEKLLEEEQAP	61		
Qy	SOOHAKVYBMRCTBYISTEFNRGINSNEKPEVKIGVSVKQOFTEEELYKRDQS	121		
Db	SOOHAKVYBMRCTBYISTEFNRGINSNEKPEVKIGVSVKQOFTEEELYKRDQS	121		
Qy	KTFDAQKSIOSHKSPRTPVEMPVFDKWNINPQAQVIFDSPKDPTSGA	181		
Db	KTFDAQKSIOSHKSPRTPVEMPVFDKWNINPQAQVIFDSPKDPTSGA	181		
Qy	MSQMIRGMDEENQFAYFLPVETLKRRKDQEENDYAPDDVYDKTAREY	241		
Db	MSQMIRGMDEENQFAYFLPVETLKRRKDQEENDYAPDDVYDKTAREY	241		
Qy	NKASKGYEENYFFPREGDGVYYNELLTRVLSKRAKAGVQSGTNALLVVKHGD	301		
Db	NKASKGYEENYFFPREGDGVYYNELLTRVLSKRAKAGVQSGTNALLVVKHGD	301		
Qy	LEAQEARQKOLENEPEEESEEMETEKEAGSSDEQEKGSSEKEGSEDEHSQ	361		
Db	LEAQEARQKOLENEPEEESEEMETEKEAGSSDEQEKGSSEKEGSEDEHSQ	361		
Qy	EEDGRDEADSKSGIGEDESEDEARAAARDKEIFGSDADSEDDADSDDEBGRQAO	421		
Db	EEDGRDEADSKSGIGEDESEDEARAAARDKEIFGSDADSEDDADSDDEBGRQAO	421		
Qy	DSDSGNGGGQRSHRSASPPFPSGEHSQAQDGSEAADSSEADSDSD	481		
Db	DSDSGNGGGQRSHRSASPPFPSGEHSQAQDGSEAADSSEADSDSD	481		

RESULT 2
 US-10-721-553-2
 ; Sequence 2, Application US/10721553
 ; Publication No. US20050032079A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barra, Surinder K.
 ; APPLICANT: Hollingsworth, Michael A.
 ; APPLICANT: University of Nebraska Board of Regents
 ; TITLE OF INVENTION: Novel Gene That is Amplified and
 ; TITLE OF INVENTION: Overexpressed in Cancer and Methods of Use Thereof
 ; FILE REFERENCE: USMN6121
 ; CURRENT APPLICATION NUMBER: US/10/721,553

CURRENT FILING DATE: 2003-11-25
 PRIORITY APPLICATION NUMBER: US/09/647,143
 PRIOR FILING DATE: 2000-09-27
 PRIORITY APPLICATION NUMBER: PCT/US99/06633
 PRIOR FILING DATE: 1999-03-26
 PRIORITY APPLICATION NUMBER: 60/079, 649
 PRIORITY FILING DATE: 1998-03-27
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 2
 LENGTH: 531
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-721-553-2

	Query Match	Score	DB	Length	Match	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps	;
Qy	1 MAPIIQQAQEDGHRRNSHRLPERRQVYCNLSLPDIPFDPKETTYPDQNRFYQ	100.0%	2764	17	60	1 YKATSLFEKHKKHDLLTTPDGLYTIDLNPDTRIDPNVLLDPADEKLLEBIIQAPTSKKR	100.0%	2764	17	60	0	;
Db	1 MAPIIQQAQEDGHRRNSHRLPERRQVYCNLSLPDIPFDPKETTYPDQNRFYQ	100.0%	2764	17	60	1 YKATSLFEKHKKHDLLTTPDGLYTIDLNPDTRIDPNVLLDPADEKLLEBIIQAPTSKKR	100.0%	2764	17	60	0	;
Qy	61 YKATSLFEKHKKHDLLTTPDGLYTIDLNPDTRIDPNVLLDPADEKLLEBIIQAPTSKKR	100.0%	2764	17	60	121 SQHQAKYVUPWKRKTKEYSTEENRYG1SNEKPEVKIGSVKQQTTEEEYKDRQSQTATE	100.0%	2764	17	60	0	;
Db	61 YKATSLFEKHKKHDLLTTPDGLYTIDLNPDTRIDPNVLLDPADEKLLEBIIQAPTSKKR	100.0%	2764	17	60	121 SQHQAKYVUPWKRKTKEYSTEENRYG1SNEKPEVKIGSVKQQTTEEEYKDRQSQTATE	100.0%	2764	17	60	0	;
Qy	181 KTFEDAKQSISOHYSKERVTPVEMPVFPDFKWINPAQIVFDSPPKDTSGAALEM	100.0%	2764	17	60	181 KTFEDAKQSISOHYSKERVTPVEMPVFPDFKWINPAQIVFDSPPKDTSGAALEM	100.0%	2764	17	60	0	;
Db	181 KTFEDAKQSISOHYSKERVTPVEMPVFPDFKWINPAQIVFDSPPKDTSGAALEM	100.0%	2764	17	60	181 KTFEDAKQSISOHYSKERVTPVEMPVFPDFKWINPAQIVFDSPPKDTSGAALEM	100.0%	2764	17	60	0	;
Qy	241 MSQAMIRGMMDDEGNOFVAYFLPVETLKKRGRDQEEMDYAPDDVYDVKJAREYNWNYK	100.0%	2764	17	60	241 MSQAMIRGMMDDEGNOFVAYFLPVETLKKRGRDQEEMDYAPDDVYDVKJAREYNWNYK	100.0%	2764	17	60	0	;
Db	241 MSQAMIRGMMDDEGNOFVAYFLPVETLKKRGRDQEEMDYAPDDVYDVKJAREYNWNYK	100.0%	2764	17	60	241 MSQAMIRGMMDDEGNOFVAYFLPVETLKKRGRDQEEMDYAPDDVYDVKJAREYNWNYK	100.0%	2764	17	60	0	;
Qy	301 NKASKGYBENYFFIFRGDGYYNELLTRVLRLSKRRAKAGVOSGTNALLVXKHDRMNEKE	100.0%	2764	17	60	301 NKASKGYBENYFFIFRGDGYYNELLTRVLRLSKRRAKAGVOSGTNALLVXKHDRMNEKE	100.0%	2764	17	60	0	;
Db	301 NKASKGYBENYFFIFRGDGYYNELLTRVLRLSKRRAKAGVOSGTNALLVXKHDRMNEKE	100.0%	2764	17	60	301 NKASKGYBENYFFIFRGDGYYNELLTRVLRLSKRRAKAGVOSGTNALLVXKHDRMNEKE	100.0%	2764	17	60	0	;
Qy	361 LEAQEARKALENHEPEEEEEEETEKEAGGSDEOEKGSSSEKEESSEDENSSESER	100.0%	2764	17	60	361 LEAQEARKALENHEPEEEEEEETEKEAGGSDEOEKGSSSEKEESSEDENSSESER	100.0%	2764	17	60	0	;
Db	361 LEAQEARKALENHEPEEEEEEETEKEAGGSDEOEKGSSSEKEESSEDENSSESER	100.0%	2764	17	60	361 LEAQEARKALENHEPEEEEEEETEKEAGGSDEOEKGSSSEKEESSEDENSSESER	100.0%	2764	17	60	0	;
Qy	421 EFGDRDEASDGSGEDESSEDBARARDKEETFGSDADSEDDADSDDBDRGQAQGGSND	100.0%	2764	17	60	421 EFGDRDEASDGSGEDESSEDBARARDKEETFGSDADSEDDADSDDBDRGQAQGGSND	100.0%	2764	17	60	0	;
Db	421 EFGDRDEASDGSGEDESSEDBARARDKEETFGSDADSEDDADSDDBDRGQAQGGSND	100.0%	2764	17	60	421 EFGDRDEASDGSGEDESSEDBARARDKEETFGSDADSEDDADSDDBDRGQAQGGSND	100.0%	2764	17	60	0	;
Qy	481 DSDSGNGGGQERSRSRSRSASPPSGSEHSQAQDGSRAAASDSEADSDP	100.0%	2764	17	60	481 DSDSGNGGGQERSRSRSRSASPPSGSEHSQAQDGSRAAASDSEADSDP	100.0%	2764	17	60	0	;
Db	481 DSDSGNGGGQERSRSRSRSASPPSGSEHSQAQDGSRAAASDSEADSDP	100.0%	2764	17	60	481 DSDSGNGGGQERSRSRSRSASPPSGSEHSQAQDGSRAAASDSEADSDP	100.0%	2764	17	60	0	;

RESULT 3
 US-10-450-763-50041
 Sequence 50041, Application US/10450763
 Publication No. US0050196754A1
 GENERAL INFORMATION:
 APPLICANT: Hyseq, Inc
 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 FILE REFERENCE: 790CIP3/US
 CURRENT APPLICATION NUMBER: US/10/450,763
 CURRENT FILING DATE: 2003-06-11
 PRIOR APPLICATION NUMBER: PCT/US1/08631
 PRIOR FILING DATE: 2001-03-30
 PRIOR APPLICATION NUMBER: 09/540,217
 PRIOR FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 09/649,167
 PRIOR FILING DATE: 2000-08-23
 NUMBER OF SEQ ID NOS: 60736

SOFTWARE: Custom
SEQ ID NO: 5041
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (377)..(428)
OTHER INFORMATION: Neuromodulin (GAP-43) proteins domain identified by eMATRIX,
accession number BL00412D, p-value=9.633e-09, raw score of 16.54

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(553)
OTHER INFORMATION: Xaa = X or * as defined in Table 2

US-10-450-763-50041

Query Match 96 2%; Score 2658.5; DB 18; Length 553;
Best Local Similarity 95.0%; Pred. No. 4.9e-160;
Matches 515; Conservative 6; Indels 11; Gaps 1;

Qy 1 MAPIOTQAEDGRPNSHTRLPERSGVYCRVKYCNSLDPIDPPDKFITYPPFDQNRFVQ 60
Db 12 MAPIOTQAEDGRPNSHTRLPERSGVYCRVKYCNSLDPIDPPDKFITYPPFDQNRFVQ 71
Qy 61 YKATSLKEOKHHDILTEPDLGVTIDLINPDYRIDPNVLLPADKLLEB1QAPTSSKR 120
Db 72 YKATSLKEOKHHDILTEPDLGVTIDLINPDYRIDPNVLLPADKLLEB1QAPTSSKR 131
Qy 121 SQHQAKVVPWNRKTEV1STEFNRGYISNEKEPEVKIGVSQKQFTBEETLYKDRDQITATE 180
Db 132 SQHQAKVVPWNRKTEV1STEFNRGYISNEKEPEVKIGVSQKQFTBEETLYKDRDQITATE 191
Qy 181 KTPEDAQKS-----1SQHQYSKPRTPVEMPVDFKWINPCQV1PDSDPA 229
Db 192 KTPEDAQKS-----1SQHQYSKPRTPVEMPVDFKWINPCQV1PDSDPA 251
Qy 230 KDTGAAALEMNSQANIRGMDEEGNQFVAYFLPVETLKRRDRDQEEENDYAPDDVTDY 289
Db 252 KDTGAAALEMNSQANIRGMDEEGNQFVAYFLPVETLKRRDRDQEEENDYAPDDVTDY 311
Qy 290 KIAREYNWNTKNSKGYYENYFF1FREGDGVYNELETVRSLKRRAKGVQSGTNALL 349
Db 312 KIAREYNWNTKNSKGYYENYFF1FREGDGVYNELETVRSLKRRAKGVQSGTNALL 371
Qy 350 VVKHRDNNEKELEAQARKAQLENHEPEEESEENETEKEAGGSDEEQKGSSSEKES 409
Db 372 VVKHRDNNEKELEAQARKAQLENHEPEEESEENETEKEAGGSDEEQKGSSSEKES 431
Qy 410 EDEHSGSESEREGDGEDEASDKSGSDEDESSDEARAARDKE1FGSDADSEDADSDDE 469
Db 432 EDEHSGSESEREGDGEDEASDKSGSDEDESSDEARAARDKE1FGSDADSEDADSDDE 491
Qy 470 DRGQAQGSNDSDSGSNGGGRSRSHRSRASPFPGSSEHSAQEDGSEAASDSSEADSD 529
Db 492 DRGQAQGSNDSDSGSNGGGRTSRHSRASPFPGSSEHSAQENGSEAASDSSEADSD 551
Qy 530 SD 531
Db 552 SD 553

RESULT 4
US-09-986-480-410
; Sequence 410, Application US/09986480
; Publication No. US2003002799A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 410
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match 89.1%; Score 2464; DB 10; Length 473;
Best Local Similarity 99.8%; Pred. No. 8.3e-148;
Matches 472; Conservative 0; Indels 0; Gaps 0;
Qy 1 MAPIOTQAEDGRPNSHTRLPERSGVYCRVKYCNSLDPIDPPDKFITYPPFDQNRFVQ 60
Db 61 YKATSLKEOKHHDILTEPDLGVTIDLINPDYRIDPNVLLPADKLLEB1QAPTSSKR 120
Db 61 YKATSLKEOKHHDILTEPDLGVTIDLINPDYRIDPNVLLPADKLLEB1QAPTSSKR 120
Qy 121 SQHQAKVVPWNRKTEV1STEFNRGYISNEKEPEVKIGVSQKQFTBEETLYKDRDQITATE 180
Db 121 SQHQAKVVPWNRKTEV1STEFNRGYISNEKEPEVKIGVSQKQFTBEETLYKDRDQITATE 180
Qy 181 KTPEDAQKS-----1SQHQYSKPRTPVEMPVDFKWINPCQV1PDSDPA 229
Db 181 KTPEDAQKS-----1SQHQYSKPRTPVEMPVDFKWINPCQV1PDSDPA 251
Qy 301 NKASKGYEEVNPF1FREGDGVYNELETVRSLKRRAKGVQSGTNALLVVKHRDNEKE 360
Db 301 NKASKGYEEVNPF1FREGDGVYNELETVRSLKRRAKGVQSGTNALLVVKHRDNEKE 360
Qy 361 LEAPARKQOLENEPEEESEENETEKEAGGSDEEQKGSSSEKES 420
Db 361 LEAPARKQOLENEPEEESEENETEKEAGGSDEEQKGSSSEKES 420
Qy 421 EEGDRDEASDKSGSDEDESSDEARAARDKE1FGSDADSEDADSDDE 473
Db 421 EEGDRDEASDKSGSDEDESSDEARAARDKE1FGSDADSEDADSDDE 473

RESULT 5
US-10-863-332-410
; Sequence 410, Application US/10863332
; Publication No. US2005006458A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: PCT/US00/12788
; PRIOR APPLICATION NUMBER: 2001-11-08
; CURRENT FILING DATE: 2004-06-09
; PRIOR APPLICATION NUMBER: US/09/985,480
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 410
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens

FEATURE: SITE
 NAME/KEY: SITE
 LOCATION: (405)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Query Match Score: 2464; DB: 17; Length: 473;
 Best Local Similarity: 99.8%; Pred. No.: 8..3e-148;
 Matches: 472; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

Qy 1 MAPTQTAQREDGRPNSHRTLPERSGYCRVYCNSLDPIDFPKFTYPFDQNRFVQ 60
 Db 1 MAPTQTAQREDGRPNSHRTLPERSGVCRVYCNSLDPIDFPKFTYPFDQNRFVQ 60

Qy 61 YKATSLKEKHKHDLTEDPIGVTIDLINPDYRIPDNVTLDADEKLLFEELIQAPTSKKR 120
 Db 61 YKATSLKEKHKHDLTEDPIGVTIDLINPDYRIPDNVTLDADEKLLFEELIQAPTSKKR 120

RESULT 7
 US-10-424-599-223174
 Qy 121 SQQAKVPMNKRTEYISTEPNRYGISNEKEPVIGSVKQQTPEEVKDRSQTAAE 180
 Db 121 SQQAKVPMNKRTEYISTEPNRYGISNEKEPVIGSVKQQTPEEVKDRSQTAAE 180

Qy 181 KTFEDAQKSISQYBKPRTTPVEWMPVFPDKWVINPCAQVIFSDPAKDTSQAALM 240
 Db 181 KTFEDAQKSISQYBKPRTTPVEWMPVFPDKWVINPCAQVIFSDPAKDTSQAALM 240

Qy 241 MSQAMTRGMDEEGHOFVAYFLPVETTLKRKDRQDDEEMDYAPDVYXYYKIAREYNNWVK 300
 Db 241 MSQAMTRGMDEEGHOFVAYFLPVETTLKRKDRQDDEEMDYAPDVYXYYKIAREYNNWVK 300

Qy 301 NKASGYENYFTIPEGGVYNELETYRLSKRAKGAVQSTGNALLVKGDRMNEKE 360
 Db 301 NKASGYENYFTIPEGGVYNELETYRLSKRAKGAVQSTGNALLVKGDRMNEKE 360

Qy 361 LEQAFARKAQLENHEPEEEEMETEEKEAGGSDEEQKGSSSEKEGSDEHSGSER 420
 Db 361 LEQAFARKAQLENHEPEEEEMETEEKEAGGSDEEQKGSSSEKEGSDEHSGSER 420

Qy 4221 EEGDRDEASDKSGSGEDDESEDDEAARDKEEKFSDASEDDEDSDDDDRGQ 473
 Db 4221 EEGDRDEASDKSGSGEDDESEDDEAARDKEEKFSDASEDDEDSDDDDRGQ 473

RESULT 6
 US-10-450-763-50040
 Sequence 50040, Application US/10450763
 Publication No. US20050196754A1
 GENERAL INFORMATION:
 APPLICANT: Hyseq, Inc
 TITLE OF INVENTION: NOVEL NUCLEAR ACIDS AND POLYPEPTIDES
 FILE REFERENCE: 790C1P3/US
 CURRENT APPLICATION NUMBER: US/10/450,763
 CURRENT FILING DATE: 2003-06-11
 PRIOR APPLICATION NUMBER: PCT/US01/08631
 PRIOR FILING DATE: 2001-03-03
 PRIOR APPLICATION NUMBER: 09/540,217
 PRIOR FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 09/549,167
 PRIOR FILING DATE: 2000-03-23
 NUMBER OF SEQ ID NOS: 60736
 SOFTWARE: Custom
 SEQ ID NO: 50040
 LENGTH: 133
 TYPE: PRT
 ORGANISM: Homo sapiens

US-10-450-763-50040
 Query Match Score: 22.5%; DB: 18; Length: 133;
 Best Local Similarity: 64.4%; Pred. No.: 7e-32;
 Matches: 130; Conservative: 0; Mismatches: 2; Indels: 70; Gaps: 1;

Qy 273 RDQEEMDYAPDVYDYLKAREYNNWVKNSKASKGYYENTFFIREGDGVYNELETRVL 332
 Qy 1 RDOBEEMDYAPDVYDYLKAREYNNWVKNSKASKGYYENTFFIREGDGVYNELETRVL 332

Db 333 SKRRKAGVQSGTNALLVKGDRMNEKELEQEARQALENHEPEEEEMETEEKEAG 392
 Qy 36 GSDEEQKGSSSEKEGSDEHSGSEREGDRDEASDKSGSGEDDESEDDEAARDKEE 452
 Db 51 GSYBEQKGSSSEKEGSDEHSGSEREGDRDEASDKSGSGEDDESEDDEAARDKEE 110

Qy 453 IFGSDADSEDDADSDDEDRGQA 474
 Db 111 IFGSDADSEDDADSDDEDRGQA 132

Db 1 RDOBEEMDYAPDVYDYLKAREYNNWVKNSKASKGYYENTFFIREGDGVYNELETRVL 332
 Qy 1 Sequence 223174, Application US/10424599
 Publication No. US2004031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J
 APPLICANT: Kovacic, David K
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 File Reference: 38-21(53223)B
 Current Application Number: US/10/424,599
 Current Filing Date: 2003-04-28
 Number of SEQ ID NOS: 285684
 SEQ ID NO: 233174
 LENGTH: 86
 TYPE: PRT
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_43557C.1.dep

Qy 66 LEKQHKHDLLTTEPDLGVTIDINPDYRIPDNVLIDPADEKLULEETIQAPISSKRSQOHA 125
 Db 1 LEKQHKHDLLTTEPDLGVTIDINPDYRIPDNVLIDPADEKLULEETIQAPISSKRSQOHA 60

Qy 126 KVVPWWKRTETVISTEFNYGINSNEKP 151
 Db 61 KVVPWWKRTETVISTEFNYGINSNEKP 86

RESULT 8
 US-10-424-599-213359
 Sequence 213359, Application US/10424599
 Publication No. US2004031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J
 APPLICANT: Kovacic, David K
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 File Reference: 38-21(53223)B
 Current Application Number: US/10/424,599
 Current Filing Date: 2003-04-28
 Number of SEQ ID NOS: 285684
 SEQ ID NO: 213359
 LENGTH: 571
 TYPE: PRT
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_34688C.1.dep

US-10-424-599-213359

RESULT 15
 US-10-689-082-4
 i sequence 4, Application US/10689082
 i Publication No. US2004014234A1
 GENERAL INFORMATION:
 i APPLICANT: FOSTER, Timothy
 i TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLO
 i FILE REFERENCE: P06315US04/BAS
 i CURRENT APPLICATION NUMBER: US/10-689,082
 i PRIOR APPLICATION NUMBER: 09/386,962
 i PRIOR FILING DATE: 2003-10-21
 i PRIOR APPLICATION NUMBER: 60/098,443
 i PRIOR FILING DATE: 1998-08-31
 i PRIOR APPLICATION NUMBER: 60/117,119
 i PRIOR FILING DATE: 1999-01-25
 i NUMBER OF SEQ ID NOS: 39
 i SOFTWARE: PatentIn version 3.1
 i SEQ ID NO 4
 LENGTH: 1742
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-10-689-082-4

Query Match 8.5%; Score 234.5; DB 16; Length 1742;
 Best Local Similarity 21.0%; Pred. No. 5.1e-06;
 Matches 122; Conservative 88; Mismatches 240; Indels 149; Gaps 24;

Qy 4B FITYPFDQNRFVQKATSLKEQKHDLITIPEPDLYTIDILNPDTYRIPD-----NVLDP 102
 Db 716 YVTLKDSNNRELQRITTDQSGHYQFDNLONGT--YTVEFAIPDNTYPSANNSTNDAIS 773

Qy 103 ADEKLLEEQAQPSSKRSQOHAKV----VPMMRKTEYISTEENRIGISNEKPEVKIG 156
 Db 774 DGERDGTRKVVWAKCTINNAADNMVTDTGFYLTPKCNVGYWWETNKDQODNEKGTS 833

Qy 157 VSV-----KQOFTEEEITY-----KDRDS Q 175
 Db 834 VKVTLKNGNDTIGTTDSNGKYEFTGLENQDYTIEPTPEGYTPKQNSGSDEGKDSN 893

Qy 176 ITAIEKTFEDA-QKSISOHYSKPRVTPVYMPVPDFKWNINPCAQVIFPDSDP-----227
 Db 894 GTKTKTVVDADNKTDISGTYKPKTN-----LGDY-TWBDTNAKDG1ODDSEKG1S4VK 945

Qy 228 -APKDTSGAA---ALEMMSOAMIRGMMDDEGNQFVAYLPVVBETLKRKRQDEEEMDY- 281
 Db 946 VTLKDQNGNAIGTTTDASHGYOPKGL--ENGSTVTEPFTPSGYTPTKANSQDITVDSN 1003

Qy 282 -----APDWDYD--YKIAr----BYNMWVRNK----ASKY-----307
 Db 1004 GITTGIIINGADNUITDISGTYKTPKYSVGDYVWEDTNKG1ODDNEKG1SGVKTLLDK 1063

Qy 308 -----EENMFPLFREGD-GVYVNELETRVLSKRAKAGYQSGTNALLYKHRDMN 357
 Db 1064 GN11STTTDENGKIQFDNIDSGNYIHHEKPEGMQTANSG-----ND 1109

Qy 358 EKELEAQARKA-----QLENHHEPEEEEEEEMETEEKAEGGSDEEQEKGSSSEKEGSE 410
 Db 1110 EKDADGEDEVTVTIDHDFFSDIDNGYFDDDSDSADSDSDSDSDSDSDSDSDSD 1169

Qy 411 DEHSSESESE-ECDRDEASDK-SESGBEDESSEDARARDKEEIGSDADSDSDSDSDSDSD 468
 Db 1170 DSDSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1229

Search completed: September 26, 2005, 13:59:26
 Job time : 176 secs

GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

Copyright

using sw model

OM protein - protein search,

Run on: September 26, 2005, 13:44:20 ; Search time 42 Seconds
Title: US-10-721-553-2
Perfect score: 2764
Sequence: 1 MAPTQTAQEDGHPRNSH.....QEDGSBAAADSSEADSDSD 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283416 seqs,
96216763 residues

Total number of hits satisfying chosen parameters:

283416

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79;*

1: pir;:*

2: pir;:*

3: pir;:*

4: pir;:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	645	23.3	425	2 T20261	hypothetical protein
2	335.5	12.1	547	2 C96828	unknown protein F1
3	280.5	10.1	457	2 P0233	probable DNA-directed hypothetical protein
4	253	9.2	445	2 S44541	neurofilament triplet hypothetical protein
5	245.5	8.9	784	2 PN0009	hypothetical protein
6	241	8.7	792	2 T4963	glutamic acid-rich hypothetical protein
7	236	8.5	590	2 A40437	hypothetical protein
8	234	8.5	953	2 C89824	hypothetical protein
9	230.5	8.3	1385	2 D89824	Clumping factor B
10	230	8.3	877	2 F90072	fibrinogen-binding protein
11	225.5	8.2	933	2 S41539	hypothetical protein
12	224	8.1	797	2 A36811	fibrinogen-binding protein
13	223.5	8.1	1315	2 T28679	probable secreted acidic repetitive protein
14	221	8.0	1192	2 A71623	hypothetical protein
15	220	8.0	334	2 A54138	ribosomal transcript
16	218	7.9	406	2 S38170	microtubule-associated protein - yeast
17	218	7.9	1166	2 T28680	glutamic acid-rich fibrinogen-binding protein
18	216	7.8	913	2 T52485	neurofilament protein NF-180 - sea lamprey
19	216	7.8	11116	2 T15116	hypothetical protein
20	214.5	7.8	1141	2 R89824	ribosomal transcript
21	214	7.7	727	2 JCS113	microtubule-associated protein - yeast
22	214	7.7	2364	2 A56577	glutamic acid-rich fibrinogen-binding protein
23	213.5	7.7	678	2 A54514	neurofilament protein NF-180 - sea lamprey
24	211.5	7.7	1092	2 T30214	hypothetical protein
25	211.5	7.7	665	2 BT1609	fibrinogen-binding protein
26	211	7.6	989	2 D89852	neurofilament protein NF-180 - sea lamprey
27	208.5	7.5	798	2 S22314	transcription factor
28	208	7.5	765	2 S70358	centromere protein
29	203.5	7.4	606	2 S70358	centromere protein

ribosomal transcri					
autoantigen NOR-90					
mature parasite-in					
immediate-early pr					
centromere protein					
transcription fact					
UBP transcription					
nucleolar protein					
transcription fact					
hypothetical prote					
TRP-containing/SH2					
microtubule-associ					
mannosid phosphoryl					
UBP transcription					
histidine-rich cal					
cyclin I - human					

30 202.5 7.3 764 2 JC5112
31 202 7.3 727 2 S18193
32 202 7.3 1526 2 A45605
33 200.5 7.3 407 1 EDBQ3
34 199.5 7.2 599 2 S18735
35 198.5 7.2 524 2 S35551
36 198 7.2 727 2 B40439
37 195.5 7.1 500 2 S55785
38 195 7.1 764 2 S0318
39 194 7.0 409 2 EB8336
40 193 7.0 1173 2 T42719
41 194.5 7.0 2464 1 QRNSP1
42 192 6.9 1178 2 S78475
43 191.5 6.9 2 A40379
44 191.5 6.9 852 2 A34373
45 191 6.9 598 2 B40713

ALIGNMENTS

ribosomal transcri					
autoantigen NOR-90					
mature parasite-in					
immediate-early pr					
centromere protein					
transcription fact					
UBP transcription					
nucleolar protein					
transcription fact					
hypothetical prote					
TRP-containing/SH2					
microtubule-associ					
mannosid phosphoryl					
UBP transcription					
histidine-rich cal					
cyclin I - human					

R.Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: T20561
A;Status: preliminary;
A;Molecule type: DNA;
A;Residues: 1-425 <WII>
A;Cross-references: UNIPROT:P90783; EMBL:Z81051; PMID:CAE02869.1;
A;Experimental source: Clone C55A6
C;Genetics:
A;Gene: CBSP:C55A6.9
A;Map position: 5
A;Introns: 14/2; 48/2; 90/3; 177/3; 381/1
RESULT 1
T20261
hypothetical protein C55A6.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Dte: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20561
R.Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: T20561
A;Status: preliminary;
A;Molecule type: DNA;
A;Residues: 1-425 <WII>
A;Cross-references: UNIPROT:P90783; EMBL:Z81051; PMID:CAE02869.1;
A;Experimental source: Clone C55A6
C;Genetics:
A;Gene: CBSP:C55A6.9
A;Map position: 5
A;Introns: 14/2; 48/2; 90/3; 177/3; 381/1
Query Match 23.3%; Score 645; DB 2; Length 425;
Best Local Similarity 33.1%; Prod. No. 3.8e-26;
Matches 146; Conservative 96; Mismatches 165; Indels 34; Gaps 10;

QY	439 SSEDEARAARDKEELFGSDAD 459 : ; : ; : ; : ; : ;	T50233 probable DNA-directed RNA polymerase II regulator [imported] - fission yeast (Schizosaccharomyces pombe)
Db	407 SDDDD--SPKKKEPVVDSQD 425	C;Species: Schizosaccharomyces pombe
RESULT 2		C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C95828	unknown protein F19K16.29 [Imported] - Arabidopsis thaliana	C;Accession: T50233
C;Species: Arabidopsis thaliana (mouse-ear cress)		R;Cadiou, E.; Lelaure, V.; Galibert, P.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.
C;Accession: C96828	#sequence_revision 02-Mar-2001 #text_change 09-Jul-2004	submitted to the EMBL Data Library, January 1999
R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anen, N.F.; Hughes, B.; Hultzar, L.	A;Reference number: Z25048	A;Accession: T50233
Nature 408, 816-820, 2000	A;Status: preliminary; translated from GB/EMBL/DBJ	A;Molecule type: DNA
A;Authors: Hunter, J.L.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroo, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	A;Residues: 1-457 <CAD>	A;Cross-references: UNIPROT:Q9US06; EMBL:AL136235; PIDN:CAB55804.1; GSPDB:GN00066; SPDB
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	A;Experimental source: strain 972h(-); cosmid c664	A;Genetics:
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.	A;Gene: SPAD664.03	A;Map position: 1
A;Reference number: A86141; PMID:11130712	A;Introns: 1/3	A;Map position: 1
A;Accession: C96828	Query Match 12.1%; Score 335.5; DB 2; Length 547; Best Local Similarity 24.8%; Pred. NO. 3.4e-10; Mismatches 69; Indels 136; Gaps 20; Matches 121; Conservative 99; Gaps 20;	Query Match 10.1%; Score 280.5; DB 2; Length 457; Best Local Similarity 24.9%; Pred. NO. 1.8e-07; Mismatches 94; Indels 61; Gaps 21; Matches 122; Conservative 94; Gaps 21;
A;Status: preliminary	QY 6 QTQIQRRED-GHR---PNSHRT-----LPERSGVVCRVKYCNC 37	QY 26 RSGVVCRVKYCNCNSLPDIPDPKFTYPPDQNRFVQYK---ATSLEKOKHKLDTLPDLG 81
A;Molecule type: DNA	Db 143 ELPEKORODKEKRHQMKNSKQMPKGHTBEKPKTPLLTDRVENRLKXKPTFLCKLKFRN 202	Db 6 RQDYILRVYRINPBLPPPPPKLINTP---NPVKQVALPNFVSTLYQEKCAPIENDIBLG 62
A;Residues: 1-547 <STO>	QY 38 SLPPDLPFDPRIFT----YFDONRFTVQYKATSYKHDLLTEPGLVTIDLAFDT 91	Db 82 VTDL-----INPDYTRID---PNVTLDPABKLLEEEIQAFTSSRSQCAKTVPMR 132
A;Cross-references: UNIPROT:Q9CA82; GB:AE005173; NID:96453869; PIDN:AAF09053.1; GSPDB:GN	Db 203 ELPPDPSAQKLMTKRDKHYDFDTRFTVYTTSLEKWKPKIFVPEPDGLIPDLDLSV 262	Db 63 MPDLLAGITGFEGDSSWMHSDSLSSNLDPIPDRSLIK----VAGGGSTHLE-VPFLLR 115
A;Gene: F19K16.29	QY 92 YRIDPNVL--LDPDAEKLUEEIQIAFTSSRSQCAKTVPM-----WNRKTEYSTFNR 143	QY 133 KTEYISTSTEENRGISNEKPEYKIGVSYKQOFTEE-EIYKORDSQTIAETKTFEDAQKSI 190
A;Map position: 1	Db 263 YN-PPRKVAPLAPPDEELLRD-DAVTPKTKRGRKERTPDKGMWVKTQISS--- 316	Db 116 RTBYISSEVAR--AASRGNRNLRTASTSKALAEQRGRSLREPKVPLAINBFDVWQPL 173
QY 144 YGISNEKPKVKIGVSYKQOFTEEII-----YKDRDSQTIAETKTFEDAQ 188	QY 191 SQ--HYSKPRVTPVEMVPVFDEKQVIFSDPAPKDTSGA----- 236	QY 191 SQ--HYSKPRVTPVEMVPVFDEKQVIFSDPAPKDTSGA----- 236
Db 3117 --INNE-----SAROSLTQEAKLREMKGGINLHNLRNRQIKDIBASFE ACK 365	Db 174 EQLKHPKTPKDLKPKPSAWNLLPNTMSAQIOQHMLRVDADDLSRSRSHSSVSSLVNLQEGHNLTK 233	Db 174 EQLKHPKTPKDLKPKPSAWNLLPNTMSAQIOQHMLRVDADDLSRSRSHSSVSSLVNLQEGHNLTK 233
QY 189 SISDHYSKPRVTPVEMVPVFDEKQVIFSDPAPKDTSGA 236	QY 237 ---ALERMQSAMIRGMDEENQTYFLPEETIKRKRQDEEMNDYAFDVKYKIA 292	QY 237 ---ALERMQSAMIRGMDEENQTYFLPEETIKRKRQDEEMNDYAFDVKYKIA 292
Db 366 SRYPATHANQLOPQEVLPILPYDRYDQEVVANFDGAPIADSEFFGKLPSTRDAHESR 425	Db 234 RHEBVALFMPSA-----EGBEPFLSYYLSSBTAE---EIQAKYNDASADVHFEPVY 281	Db 234 RHEBVALFMPSA-----EGBEPFLSYYLSSBTAE---EIQAKYNDASADVHFEPVY 281
QY 237 ALERMQSAMIRGMDEENQTYFLPEETIKRKRQDEEMNDYAFDVKYKIAEYN 296	QY 293 REY-NWNVVKNAKSKGIGBENYFFIFR-----EGDGYYNELETRVLSKRRAKAGVOSGT 345	QY 293 REY-NWNVVKNAKSKGIGBENYFFIFR-----EGDGYYNELETRVLSKRRAKAGVOSGT 345
Db 426 AI--LKSYYVAGSDIANPFLATMPSDLSIDHNEEISYT-----WVEYL 475	Db 282 NFNFPTDASHMINSTGLCLTPTFDKDHRPANQOLYTPARSTSKRHTRAVPSLDA 341	Db 282 NFNFPTDASHMINSTGLCLTPTFDKDHRPANQOLYTPARSTSKRHTRAVPSLDA 341
QY 297 WNWVVKNAKSKGIGBENYFFIFR-----VSFDNTASYLVYSSR- 505	QY 346 NALLYVCHRDNEKE-LEAQEAR--KAOLHENHEPEEEEFEMETEEKEAGGSDEEEQKGS 402	QY 346 NALLYVCHRDNEKE-LEAQEAR--KAOLHENHEPEEEEFEMETEEKEAGGSDEEEQKGS 402
Db 476 WDQPNAN-----DPCTYL----- 505	Db 342 VDGIELLRLDNEESEQKLPKARYDFGFLGNIKDLEEEEEKRSVSE--GSUNEE---L 394	Db 342 VDGIELLRLDNEESEQKLPKARYDFGFLGNIKDLEEEEEKRSVSE--GSUNEE---L 394
QY 357 NEKLEAQARKADQEN-----HEEEEEEMETEEKEAGGSDEEQEKSSE 407	QY 403 SSEKEGSEDEHGSESSREEGDRDESDKSSEGEDESSEDARA-ARDKEEFLGSDADSE 461	QY 403 SSEKEGSEDEHGSESSREEGDRDESDKSSEGEDESSEDARA-ARDKEEFLGSDADSE 461
Db 506 ---VGASSSKMRLEDGGLGRSSWKHEFQD----- 539	Db 395 SBEBKPAESREQUESEAQTNGYKTPETOQAQNM-----ASESQANSAPPVEE-GNTQPSP 448	Db 395 SBEBKPAESREQUESEAQTNGYKTPETOQAQNM-----ASESQANSAPPVEE-GNTQPSP 448
QY 408 GSDEDEHS 414	QY 462 DDADSDDD 470	QY 462 DDADSDDD 470
Db 540 GNEDDS 546	Db 449 VEQLQNEED 457	Db 449 VEQLQNEED 457

RESULT 3

A;Status: translation not shown

A;Accession: S44541

A;Title: The sequence of a 32120 bp segment located on the right arm of chromosome II f

A;Reference number: S44537; PMID:94378722; PMID:8091861

A;Map position: 1

A;Introns: 1/3

RESULT 5

A; Molecule type: DNA
A; Residues: 1-45 <SH>
A; Cross-references: EMBL:236148; PID:9536721; PTN:CA85243.1; PID:YBR279w
R; Brandt, T.; Christiansen, C.; Holmstrom, K.; Kallehoe, T.
submitted to the Protein Sequence Database, August 1994
A; Reference number: S46157
A; Accession: S46161
A; Molecule type: DNA
A; Residues: 1-45 <DNA>
A; Cross-references: EMBL:236148; PID:9536721; PTN:CA85243.1; PID:YBR279w
R; Shi, X.; Finkelstein, A.; Wolf, A.J.; Wade, P.A.; Burton, Z.F.; Jaehning, J.A.
Mol. Cell. Biol. 16, 669-676, 1996
A; Title: Paf1p, an RNA polymerase II-associated factor in *Saccharomyces cerevisiae*, may
A; Reference number: JC6088; PMID:9614034; PMID:8552005
A; Accession: JC6088
A; Molecule type: DNA
A; Residues: 1-166-168-445 <SH>
A; Experimental source: strain YJJ453
A; Accession: PE6031
A; Molecule type: DNA
A; Residues: 5-11-420-427 <SH>
C; Comment: This factor is a highly charged nuclear protein, and acts as a cofactor important for transcriptional regulation.
A; Gene: SGD:PAF1
A; Cross-references: SGD:S0000483; MIPS:YBR279w
A; Superfamily: Saccharomyces cerevisiae hypothetical protein YBR279w
A; Map position: 2R
A; Note: this gene is located at the right arm of chromosome II
C; Superfamily: *Saccharomyces cerevisiae* hypothetical protein YBR279w
C; Keywords: nucleus
F; 25-49/Region: PEST sequence
F; 119-141/Region: nuclear location signal

Query Match Score 9.2%; Best Local Similarity 22.1%; Pred. No. 4.4e-06; Mismatches 96; Indels 184; Gaps 21;

Query 23 LPERSSEVCRVKYCNSLPDPIDPDKPITYP-----FDQNRFVQYKATSLEKOHK 71
Db 1 MSKQPYIYAPKIQNLPLPPLPKLVLVPPSPETNAQQSQLINSLYIKNNVNLIQQ -- 58

Query 72 HDLITEPDLYTIDI-----NPDTYRIDPNVLLDPADEKCLLEBHQAPTSSKRS 121
Db 59 -----DEDGLMPVDLJKFPGHLNKLDLSKLJXGFD-NVKLKDRLRDRD-----PRIDRIT 108

Query 122 QOHAKVVPWKRKTETISTEPNRGTSNEKEPVKIGVSVKQFTEBIVKDRDSQ ----- 175

Db 109 KTDISKVTFARTEFTSNTAAHDNTSLKRRKL-----DDGDSDENLDV 154

Query 176 ---ITAIEKTFEDADLSISCHYHSKPRVTPEVMPVFDPFMWNPCAQT/FDSDPAPKDT 232

Db 155 NHISAVEGTFNKTDP---WOPPVKGKvWvKWDLDP---TASMDQVTF----ILKF 203

Query 233 SGAAALEMMQAMIR GM--MDEEGNQFYAYPLVEETIKKKRDKQEEEMDYAPDDVYD 288
Db 204 MGSAISLDTKERKSLSLTGIFRPELEEDENISMYATDHKDOSAILENLEKOMDENDDSHE 263

Query 289 ---YKLAETNWVKNKAQGKSYEENYFFIREGDGV-YINBLETRVLSKRAKAGVQSG 344

Db 264 GKIVKFKRIRDYDMQVAEFPMTE-LAIRNDKGIAYYKPLRSKIELARRVNDLIIKP- 321

Query 345 TNALLYVKH-----ROMNEKELEAQAFARKAQLEN-----HEPEREEEFMETEKK 389

Db 322 ---LYREHDDQLNNTLNPSTKEANIRKLRMKFDPINPATVDEBDDDEBOPDVYK 377

Query 390 EGGSDEEQKGSSSEKEGGSEDEHSGSESESREGDRDEASDKSGSGEDESSEDARAARD 449
Db 378 ESEG--DSKTEGSEDEGENSKDEETKQEKNEQ-----DEENKODENRAADT 422

Query 450 KEEIFGSDAEDDALSDDDRGQAOQ 475

Db 423 PET--SDAVHTEQPSEEKETLQEE 445

RESULT 5

PN0009 neurofilament triplet M protein - Pacific electric ray (fragment)
C;Species: Torpedo californica (Pacific electric ray)
C;Date: 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: PN0009
R;Inrial, M.; Scheiller, R.H.
J. Neurochem. 54, 762-770, 1990
A;Title: A unique neurofilament from *Torpedo* electric lobe: sequence, expression, and localization
A;Reference number: PN0009; PMID:90155300; PMID:2106008
A;Accession: PN0009
A;Molecule type: mrNA
A;Residues: 1-784 <LIN>
A;Cross-references: UNIPROT:Q7LZ90
C;Comment: Neurofilaments are a subgroup of intermediate filaments which are expressed in all peripheral nerve fibers.
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; cytoskeleton; intermediate filament; nerve; phosphoprotein; tar
F; 1-52/Region: carboxyl-terminal #status predicted <CTD>
F; 53-84/Region: serine-rich
F; 616-622,628,634,640,646,652,658,670/Binding site: phosphate (Ser) (covalent) #status F
F; 98-194/Region: coil Ia
F; 217-367/Region: coil II
F; 59-59/Region: glutamic acid-rich
F; 598-674/Region: f-residue repeats
F; 675-784/Region: carboxyl-terminal #status predicted <CTD>
F; 616,622,628,634,640,646,652,658,670/Binding site: phosphate (Ser) (covalent) #status F
Query Match Score 8.9%; Best Local Similarity 20.5%; Pred. No. 2.1e-05; Mismatches 102; Indels 143; Gaps 23;

Query 47 KFITYPFDNRFVQYKATSLEKQ--HKGDLLTBPDLG-----VTDILINPDYR 93
Db 62 RFASY-IDKPHYLQSQNKELEAQHAKQVSHGQLGDVYDQEIRELRSIQVNQKAAQ 120
Query 94 ID-PNVLL-----PADEKLLEE-----IQAFTSSKESQHQAKVY 128
Db 121 IQLDVSHLDDDFQRVGAFDEALRDEDPBEATIRVLKQETEDSVIQAGDGEKHAOSLQDEV 180
Query 129 PWRKTTVISTEPNRGTSNEKEPVKIGVSVKQFTEBIVKDRDSQTAIKTFEDAKQ 188
Db 181 AFLR-----NNHEEV-AIDLFAQTOATQTVVEK-KDFLFKTDITSALEKEIRS 224
Query 189 SISQHYSKPRVTPEVMPVFDPFKMW-----INPACQIVFDDPAPKDTGAAALENM 241
Db 225 QLBESHKAKMQQADE-----WPKCRYDKLINEALEMNKAIRAAREEIGYRQLQ 274
Query 242 SQAM---IRGM-----MDEEGNQFYAYPLVPEETLKCKRKDQEEN----DYA 282
Db 275 SKSIBLESPTRSKESLRLQTLDEDRIHRADVANTQETVQOLENLRTGTKENAHRLEY- 333
Query 283 PDDVYDKTARETNWVKNKASGYENYFFIREGDGVYNN----ELETIVRLSKRKA 337
Db 334 -QDILNWTMVALDIEIAVRLKJLGEESERYTF-SGTGPSTPYRSRSPRPLPAVTKTKEVP 391
Query 338 KAGVQSTGNALLVVKHRDNKNEK---TELAQEAKRKAQLENHEPPEEEBBMPEEE- 388
Db 392 KVVKQHKFVTEELIETKYKDEAKMGIIDLAEEVGGATMTESFDEKEAKVYEEAVAT 451
Query 389 -----KRGGSDEEQKGSSSEKEGGSEDEHSGSESESREGDRDEASDKSGSGEDESSD 442
Db 452 VQAGVQAEPRGEAESEKAESEKEDEGVBEEECKE-EADDEEKEDEBGAEDAEDEAEGG 510
Query 443 BARARDKEIF-----GSDADSEDDADSDDDEGQAOQGSNDNSDGSNGGGQR- 492
Db 511 ESRVVEKVEIVKVEQSKAHPGKDEVEKERKEEEEEESEASGESDKESTGKALINGSQEE 570
Query 493 -----SRSHSRSASPFSSEHSAQED---GSEAADSSEADSOSD 530
Db 571 SKGKVVEKLTVTEKATEDKVSPIREEKPKQEKDIEEKCEAKSKDEAKSKDEA 626

Qy	145 G ISNEKPEVKIGVSKQOFTEEIIYKORDSQITAIKTFEDAQKSISQHWSKPRVTPVET 204	Db	935 DENDKVLKTVTTDENGKYQFTDLNNGTYKVBFETPSGYPT-----SVTSGN 981		
Db	476 SNSNSYSHVN-GSSTAN-GDKQKKNLGD-----YWEWTNKDKQDANEKG1KGVYV 525	Qy	227 PAPDTSGAAALEMNSQAMTRGMIDBEG-----NQTVAYFLEVEETLKRKRDQEE 277		
Qy	205 MPVFPDFRMWINPQAOVIFDSDAPKDTGAAALEMNSQAMTRGMIDBEG-QF----- 257	Db	982 DTEKDNGLTTTGVIKDA-DNMFLDGSFVYKTPKSYLGDYWW----DSMKDGKDSTE 1034		
Db	526 I-----LXDSNG-----KELDRTTTIDENGKYQFTGLSNG 554	Qy	278 EMDYAPPDVYDYKI-----AREYWNVNKRASKYEEENTFYFIFRSQDGYYNEL 326		
Qy	258 ---VAYPLPVEETLKKRDKQEEEMDY---APFDYD---YKIAI---EYN 296	Db	1035 K-----GIKDVKVVILNEKGVEIGTITKDENGKRFDFNLDGSKYKVFEPKTGL---- 1083		
Db	555 TYSVEFSTPAGYFTTANAGTDADSDGTLTTGVIKDAADNTLDSGFYTKPSLGDYV 614	Qy	327 ETRYVLSKRRAKAGVQYSGTNALLVYKHDNMKELEAQEARAKAQNEN-HREBEEEEEEME 385		
Qy	297 WNWYKIK-----ASK3Y-----BENYFFIFREGDGVYYNELETRY 330	Db	1084 -----TGTGTTNTTEDRDADEGDEVDTIDDDFTLNDGTYESETSSDSD 1129		
Db	615 WYDSNKDGKODSTHEKG1KGVKVTLQNEKEKGVEIGTITDENGKRYFEDNLDSGKY----KV 669	Qy	386 TEEKEAGGSDEEQKGSSEKEGSDEEQKGSSESERE-EGGRDEASDKSGSGDESEDEA 444		
Qy	331 RLSKRRAKAGV-QSGTNALLVYKHDNMKELEAQEARAKAQNENHPEEE---EEBEMET 386	Db	1130 SDS 1188		
Db	670 IFEK---PAGLQTGTNTTEDDKDADGEDEVDTIDDDFTLNDGTYEEBETSDSDSDS 726	Qy	445 RAARDKBEIIFGSDADSEDDADSEDDDGRQAQGGSDNDSDGSNGGGQRSSHSRSASPP 504		
Qy	387 EKEAGGSDEEQKGSSEKESEDEHSSEBERE-TSGSEDDESED 442	Db	1189 DSD 1247		
Db	727 SDS 786	Qy	505 SGSEHSAQEDGSEAAASDS-SEADSDS 531		
Qy	443 EARAARDKEIEFGSDADSEDDADSDDEDRGQAQGGSDNDSDGSNGGGQRSSHSRSASP 502	Db	1248 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1275		
Db	787 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 845	RESULT 10			
Qy	503 FPGSEHSAQEDGSEAAASDS-SEADS 531	F90070			
Db	846 SDSDSDSDSDSDSDSDSDSDSDSDS 875	Clumping factor B [imported] - Staphylococcus aureus (strain N315)			
C;Species: Staphylococcus aureus					
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004					
C;Accession: F90070					
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.					
Lancet 357, 1225-1240, 2001					
A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.					
A;Reference number: A89758; PMID:21311952; PMID:11418146					
C;Status: preliminary					
A;Molecule type: DNA					
A;Cross-references: UNIPROT:Q99W47; GB:BA000018; PID:913700454; PIDN:BAB41751.1; GSDB:G					
C;Generics:					
A;Gene: sdrD					
Query Match Score 230.5; DB 2; Length 1385;					
Best Local Similarity 20.4%; Pred. No. 0.00023; Mismatches 89; Indels 127; Gaps 20;					
Matches 116; Conservative 89; Mismatches 236; Indels 127; Gaps 20;					
Qy	27 SGVVCRVVKYCNSLFDIIPPDKPFKTTYPDQNRREVQYKATSLSKQKHDLTTEPDLYTIDL 86	Db	318 NINGQFSLPLFTDRAKPSGTYDANIINI-ADENNNNKITYNNSSPIAGIDKPNGANANIS 375		
Db	772 TGVI-----NGADNMFLDGSF--YKTPKYNLGNYYWEDTMKDGKDSTERGISGTVYL 823	Qy	122 QQHAKVVPWNRKTEVISTEPNRYG1SNERPEVKIGVSUQQTETEYKORDSQTIAIE- 180		
Qy	87 INPD-----PRVLLDPADEKLLEBIQAP 115	Db	376 SQIGVDTASGQNTYKQTVF-----VNPKQVRVLGNTWVYIKGYQKTI-EBSGKRVSATDT 429		
Db	824 KNENGEVLTQTTKTDKGKYQFTGLENTGYKEFETPSGYTPQVGSGTDEG-IDSGNTST 882	Qy	181 --KTPF--DQKTSISQHYSKPRVPTV-EYMPVFPDKWNPQCAQVTFQDSDPAPKDTSGA 235		
Qy	116 TSSKRQSOHAKV-----VPMWRKTEVISTEPNRYG1SNEKPEVKIGVSQQTETEYK 170	Db	430 KLRIFEVNDTSKLSDSYYADPNDNSLKEVTDQFKRNIYEHPPVNAISKFD----- 480		
Db	883 TGVIKDKDNIDTSQFXTPLGTYWEDETNKNGYDODKEGISEGTV-----TLK 934	Qy	236 AALEMMSQAMTRGMIDDEGNOFQVAFYFLPVVBETLKRKRQDEEEMDYAPDYYDVKAREY 295		
Qy	171 DRDSQITAIKTFEDAOKSISQ---HYSKPRVTPVEMMPVYFPDFKWMINPCAQVIFDS 226	Db	481 --ITKTYVVLVEGHYDNTG-----KMLKTVIQENVDPVTNRDTSI----F 520		
C;Experimental source: strain N315					
C;Genetic:					
A;Accession: D89824					
A;Status: preliminary					
A;Molecule type: DNA					
A;Residues: 1-1825 <CUR>					
A;Cross-references: UNIPROT:Q99W47; GB:BA000018; PID:913700454; PIDN:BAB41751.15					
C;Generics:					
A;Gene: clfb					
Query Match Score 230.5; DB 2; Length 1385;					
Best Local Similarity 20.4%; Pred. No. 0.00023; Mismatches 89; Indels 127; Gaps 20;					
Qy	27 SGVVCRVVKYCNSLFDIIPPDKPFKTTYPDQNRREVQYKATSLSKQKHDLTTEPDLYTIDL 86	Db	318 NINGQFSLPLFTDRAKPSGTYDANIINI-ADENNNNKITYNNSSPIAGIDKPNGANANIS 375		
Db	772 TGVI-----NGADNMFLDGSF--YKTPKYNLGNYYWEDTMKDGKDSTERGISGTVYL 823	Qy	122 QQHAKVVPWNRKTEVISTEPNRYG1SNERPEVKIGVSUQQTETEYKORDSQTIAIE- 180		
Qy	87 INPD-----PRVLLDPADEKLLEBIQAP 115	Db	376 SQIGVDTASGQNTYKQTVF-----VNPKQVRVLGNTWVYIKGYQKTI-EBSGKRVSATDT 429		
Db	824 KNENGEVLTQTTKTDKGKYQFTGLENTGYKEFETPSGYTPQVGSGTDEG-IDSGNTST 882	Qy	181 --KTPF--DQKTSISQHYSKPRVPTV-EYMPVFPDKWNPQCAQVTFQDSDPAPKDTSGA 235		
Qy	116 TSSKRQSOHAKV-----VPMWRKTEVISTEPNRYG1SNEKPEVKIGVSQQTETEYK 170	Db	430 KLRIFEVNDTSKLSDSYYADPNDNSLKEVTDQFKRNIYEHPPVNAISKFD----- 480		
Db	883 TGVIKDKDNIDTSQFXTPLGTYWEDETNKNGYDODKEGISEGTV-----TLK 934	Qy	236 AALEMMSQAMTRGMIDDEGNOFQVAFYFLPVVBETLKRKRQDEEEMDYAPDYYDVKAREY 295		
Qy	171 DRDSQITAIKTFEDAOKSISQ---HYSKPRVTPVEMMPVYFPDFKWMINPCAQVIFDS 226	Db	481 --ITKTYVVLVEGHYDNTG-----KMLKTVIQENVDPVTNRDTSI----F 520		
C;Experimental source: strain N315					
C;Genetic:					

Query Match		Score	Length	DB	2;	Translated from GB/EMBL/DBJ
Best Local Similarity	21 %	Score	223.5 ;	DB	2 ;	
Local Residues	1-1315 <JOS>	Pred.	No.	0.00049 ;		
Matches	85 ;	Mismatches	225 ;	Indels	127	
Cross-references: UNIPROT:O86488 ; EMBL:AJ005646 ; NID:e1318791 ; P						
Genetics:						
Gene: sdrd .						
Y	27 SGVVERVKYCNSLPDIPIFDPKITYPFDQNRVFOVQKATSLEKOHKRHDLLTEP-	TYRID	-	PNVILDPADEKL-	-	
b	772 TGVIL---NGADNNTLDSCF--YKTPKVNGLNTWEDTNKGDKODSTEGK-	: :	:	:	:	
b	87 INPD-----	;	;	;	;	
b	824 KNEEVGVLQTTKTDQDGKYQRTGLENGTYKVEFETSGTGTQVESGSTDG-	;	;	;	;	
y	116 TSSKRQSQHAKV- ---VPAMRKTEYISTENRGYISNEKPEVKIGVSYKQQ-	;	;	;	;	
b	883 TGVIKDKDNNTIDSFGYKPTYNLGDVWEOTNKNVGQDKEBKGSQGTV- --	;	;	;	;	
b	171 DRDSQIATAEETFRDAQKSISQ - -HYSKPRVTPEVMMVFDPDFMWINPC-	;	;	;	;	
b	935 DENDKVLUKTIVTDRGKYQFTDLNNGTYKVEFETSGTGT- --	;	;	;	;	
b	227 PAPKDTSGAAALENDISOAMTRGMMDEEG- ---	;	;	;	;	
b	982 DTEKDSNLGLTTGKTKDA - -DNMTLDSFGYKPTKPSLGDYWW- ---	;	;	;	;	
b	1035 K-----GIRDVKYTL--LNEKGVEIGITKTDEKGKYCDNLDGSKY- ---	;	;	;	;	
b	336 RAKAGV - QSGTNAIIIVVKGHDMEKELEAQPARKAQLENHHEPEEEEREMET-	;	;	;	;	
b	1080 --PAGLQTGTNTTDDDKDADGEVDTITDHDFTLNGYYEBTSID- ---	;	;	;	;	
b	395 DEEQEGKGSSESEKEGSEDEEHSGSESEEREQGRDEASDKSGGEDSESEDEARA-	;	;	;	;	
b	1127 DSDDSDSDSDRSDSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD-	;	;	;	;	
b	455 GSDADSEDDADSDDDDRGQAQGSDDSDSGNSGQRSSRSASPPPSGG-	;	;	;	;	
b	1185 DSDDSD-	;	;	;	;	
b	515 GSEAAGADS-SEAUDSD 531	;	;	;	;	
b	1238 SDSDSDSDSDSDSDSDSD 1255	;	;	;	;	
RESULT 14						
71123	probable secreted protein PFB0115w - malaria parasite (Plasmodium)					
	Species: Plasmodium falciparum					
	Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09					
	Accession: A71623					
	Gardiner, M. J.; retellin, H.; Carucci, D.J.; Cummings, L.M.; Arav, Zhou, L.; Sutton, G.G.; Clayton, R.; Pertea, M.; Salberg, S.					
	Science 282, 1126-1132, 1998					
	Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum					
	Reference number: A71600 ; NUID:99021743 ; PMID:9804551					
	Accession: A71623					
	Genetics: sdrd ; static ; probable ; nucleic acid sequence not shown ; translation					
	Molecule type: DNA					
	Residues: 1-1192 <GAR>					
	Cross-references: UNIPROT:Q96127 ; GB:AE001373 ; GB:AE001362 ; NID:3D7					
	Experimental source: clone					
	Genetics: PFB0115w					
	Query Match	Score	221 ;	DB	2 ;	Length 1192

Best Local Similarity 20.2%; Pred. No. 0.00059;		Matches 100; Conservative 90; Mismatches 196; Indels 132; Gaps 20;	
Qy	55 QNRFYQKATSLR-----KOHKHDLTLEPLGLYTIDLINPDTYRIDPNVLDPADEKL 108		
Dbs	296 QSKQERQTIBLDLNGKELLKSHAN-----KIDTNTI-----ADIK 341		
Qy	109 EEEIOAPTSKRS-QQHAKVUPWMPRKTEYISTEFNRYGINSNEKPEVKIGVS/KQQFTEE 167		
Dbs	342 KEERETKDEKEKNTCQLVKDVLQLKVG-----ETKDEDEKREGTDDEDTDBE 392		
Qy	168 IYKDR----DSQIALEKTPEDAQSISQHYSKPRVTPVMPVFPDFRMWNPACQVF 223		
Dbs	393 DTDEDEEDTDEEDSDEBETGDOENKEETEYDEKCTEKA-----EBELYS 437		
Qy	224 DSDPAPKOTSQGAALAEMLMSQAMIRGMDEBGNQFVAYFLPIVEETLKERKRDDEEMDYAP 283		
Dbs	438 DKBSSEKDKREESEDKKEESE-----KDKEES-----EKDKEKTEEDDEKTEDEKG 482		
Qy	284 DDVY-----DVKAREYQNNWNVKNKASKGTYEENYFFIFREGDGVYYNELEPRVRLSKRR 316		
Dbs	483 TEVYKKETDVDEKEKGEYGEYGETDDEBKEE-----DDEETKVEECKTE 528		
Qy	337 AKAGVQSGTNALLYVKHRDMNEKELEAQ-EARKAQLENHEPEEEEEE-EHETEKEAG 392		
Dbs	529 KD----EEGTD---YEBDTDDSDKDEBTKVEBEKKTDEEEETDEKTEVEKCKTEKDEE 582		
Qy	393 GSDBEQQEKGSSSE-----KEGSBDEBHSGSE-----SEREBDRDEASD 430		
Dbs	583 GTDDEEDTDDSDKDVEETVEBETDAEDEKEENBEGTDDDEKVETEETLDDQEEQEDKEDD 642		
Qy	431 KSGGGEDESSDEAARARDKEEKFGS-----ADSEDDADSDDDERGQAGGSNDSDGSN 487		
Dbs	643 KEKDKEDDEKDKEDKEKDKEDKEKDKEDKEKDKEDKEKDKEDKEKDKEDKEKDKED 702		
Qy	488 GGGRSRSRSRSASAPPGSEHSAQGDGSEAASDSSEAASDSD 531		
Dbs	703 KEKDKEDDEKDKED-----KBDNKEKDKEDKEKDKED 734		
RESULT 15			
Qy	A54138 acidic repetitive protein ap1 - Tetrahymena thermophila C;Species: Tetrahymena thermophila C;Date: 29-Sep-1999 #sequence_revision 29-Sep-1999 #text_change 09-Jul-2004 C;Accession: A54138 R Heinonen, T.Y., Peariman, R.E. J. Biol. Chem. 269, 17428-17433, 1994 A;Title: A germline-specific sequence element in an intron in Tetrahymena thermophila A;Reference number: A54138; PMID:8021245 A;Accession: A54138 A;Molecule type: DNA A;Genetic code: SCG5 A;introns: 64/1, 158/1		
Query Match Score 8.0%; Score 220; DB 2; Length 334;			
Best Local Similarity 27.6%; Pred. No. 0.00016;			
Matches 64; Conservative 43; Mismatches 85; Indels 40; Gaps 11;			
336 RAKAVQSGTNALLYVKHRDMNE---KELEAQEARKAQLENHEPEEE-----EEE 381			
24 RKEPIQKHSIA--VSKEETENTPKLQDDEBNADEGDGIDDESGSDDDSGDDBE 81			
382 EEMETEEKEAGGSDDEOEGKSSSEKEG-SEDEHSSESESEREGDRD---BRSDKSS--- 433			
82 ESGISDDDESGSDDQESGSDDDESGSDDDESGSDDDNGSDDDNGSD 141			

Tue Sep 27 06:55:05 2005

us-10-721-553-2.rpr

Page 8

Qy 434 --SGEDSESE-- -----DEARARARDKERRIGSDADSEDDA-DSDD-BDRGORAQGSD- 479
Db 142 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 142 EDNGDDDSNDDNDGDDENGGDAEDGDDAED-GDAEDGDAEDEDGDDAEDGDDAEDGDDA 199

Qy 480 NDSDSGSGNCCQRSRSHRSASPPSGSHEHSQAQDENGSEAAASDSEASDSD 531
Db 200 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 200 EDGDDAEDGDAEDGDDAEDGDDAEDGDDAEDGDDAEDGDDAEDGDDAEDGDDAEDGDDNE 251

Search completed: September 26, 2005, 13:55:35
Job time : 45 secs

Copyright (c) 1993 - 2005	GenCore version 5.1.6	Compugen Ltd.
run on:	September 26, 2005, 13:35:20 ; Search time 167 seconds (without alignments)	QEDGSSAAASDSSEADSDSD 531
protein search model	using sw model	1229.759 Million cell updates/sec
title:	US-10-721-555-2	
target score:	2764	1 MAPTITQAFREDGHPRNSH.....
segment:	QEDGSSAAASDSSEADSDSD 531
scoring table:	BLOSUM62	
Gapext:	0.5	
seqs. searched:	2105692	386760381 residues

प्रस्तुति ३

ADO58688 ID ADO58688 standard; protein; 531 AA. XX	Qy 121 SQHAKVVPWWRKTEYISTEPNRYGISENEKPEVKIGYSVKQQFTEEYKDDDSQTIAE 180 Db 121 SQHAKVVPWWRKTEYISTEPNRYGISENEKPEVKIGYSVKQQFTEEYKDDDSQTIAE 180
AC ADO58688; XX 15-JUL-2004 (first entry)	Qy 181 KTFEDAKSISOHYSKPRVTIPVEVMPFPDFKWINCPAQIVFDPDFKAPKDTSGAALEM 240 Db 181 KTFEDAKSISOHYSKPRVTIPVEVMPFPDFKWINCPAQIVFDPDFKAPKDTSGAALEM 240
DE Human regulatory molecule HRM-9. XX cycostatic; immunomodulator; agonist; antagonist; gene therapy; human regulatory molecule; HRM; disease development; cell proliferation; immune response; cancer.	Qy 241 M\$QAMIRGMMDREGNQTYAYPLPVEETLKKRKRDKDQEEMDYAPPDVYDVKIAREYNNVVK 300 Db 241 M\$QAMIRGMMDREGNQTYAYPLPVEETLKKRKRDKDQEEMDYAPPDVYDVKIAREYNNVVK 300
OS sapiens. XX US200205264-A1. XX 16-MAY-2002. PD 26-SEP-2001; 2001US-00840787. PF 23-SEP-1997; 9705-00983750 PR 20-JAN-1999; 9905-00284613 PR 03-MAR-2000; 2000US-00538865. PA (INCYT) INCYTE PHARM INC. PI Lal P, Hillman JL, Bardman O, Shah F, Au-Young J, Yue H; PI Guegler MJ, Corley NC; XX WPI; 2004-459763/43. DR N-PSDB; ADO58737. XX PT New human regulatory molecules, useful in the diagnosis and treatment of PT cancer and immune disorders. XX PS SEQ ID NO 9; 116pp; English. XX	Qy 301 NKASKGIGENYPPIFRGDGVYNELETRVLISKRAKAGVQSGTNALLVYTHRDNEKE 360 Db 301 NKASKGIGENYPPIFRGDGVYNELETRVLISKRAKAGVQSGTNALLVYTHRDNEKE 360 Qy 361 LEQEARKAQLENHEPPEEEEMETEKEAGGSDEEQEGKSSSERGSESEDHSGSSEER 420 Db 361 LEQEARKAQLENHEPPEEEEMETEKEAGGSDEEQEGKSSSERGSESEDHSGSSEER 420 Qy 421 EFGDRDASDKSGSGGEDESSEDARARDKEIFGSDADSEDDASDDEDREOAOQGSDN 480 Db 421 EFGDRDASDKSGSGGEDESSEDARARDKEIFGSDADSEDDASDDEDREOAOQGSDN 480 Qy 481 DSDSGNSGGQRSRSHRSRASPFPGSEHSQAQDGSAAADSSEADSDSD 531 Db 481 DSDSGNSGGQRSRSHRSRASPFPGSEHSQAQDGSAAADSSEADSDSD 531
XX RESULT 4 ID ABM82102 standard; protein; 531 AA. XX AC ABM82102; XX DT 18-NOV-2004 (first entry) DE Tumour-associated antigenic target (TAT) polypeptide PR083014, SEQ: 5424. XX Tumour-associated antigenic target; TAT; human; overexpression; cancer; KW tumour; diagnosis; cell proliferative disorder; breast cancer; KW colorectal cancer; lung cancer; ovarian cancer; liver cancer KW central nervous system cancer; bladder cancer; pancreatic cancer; KW cervical cancer; leukaemia; melanoma; chromosome identification; KW chromosome mapping; gene mapping; gene mapping; KW gene therapy; cytostatic. XX Homo sapiens. OS WO2004030615-A2. XX PD 15-APR-2004. XX PF 29-SEP-2003; 2003WO-US028547. XX PR 02-OCT-2002; 2002US-0414971P. XX PA (GETH) GENENTECH INC. XX PI Wu TD, Zhang Z, Zhou Y; XX DR WPI; 2004-347921/32. XX N-PSDB; ACN40565.	
SQ Sequence 531 AA;	Query Match 100.0%; Score 2764; DB 8; Length 531; Best Local Similarity 100.0%; Pred. No 1..36-196; Indels 0; Gaps 0; Matches 531; Conservative 0; Mismatches 0;
Qy 1 MAPTQTOAQREDGERPNSHRTLPERSGVCRVYCVNSLFDIPIPDKFITYPPDNQREVQ 60 Db 1 MAPTQTOAQREDGERPNSHRTLPERSGVCRVYCVNSLFDIPIPDKFITYPPDNQREVQ 60	XX New tumor-associated antigenic target polypeptides and nucleic acids, PT useful in preparing a medicament for treating or detecting a PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or PT prostate cancer or tumor.
Qy 61 YKATSLIEQKHKHDLIIPDGLVITDILNPITYRIDPNVLLDPADSKLLEB1QAPTSSKR 120 Db 61 YKATSLIEQKHKHDLIIPDGLVITDILNPITYRIDPNVLLDPADSKLLEB1QAPTSSKR 120	XX Claim 12; SEQ ID NO 5424; 7273PP; English.
CC The invention relates to human tumour-associated antigenic target (TAT) CC polypeptides, and their related nucleic acids. The TAT polypeptides are	XX

cc overexpressed in cancer tissues compared to normal tissues, and may thus serve as effective targets for the diagnosis and treatment of cancer in mammals. The invention also relates to nucleic acid and polypeptide sequences at least 80% identical to the TAT nucleic acids and polypeptides; expression vectors and host cells comprising a TAT nucleic acid; an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAT polypeptide of the invention.

Sequence 531 AA;

Query Match Score 2764; DB 8; Length 531;
Best Local Similarity 10.0%; Pred. No. 1. 3e-196;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPTQTOAQREDGRPNSHRTLPERSGTVCRYKCNNSLPDTPFDPKPTYPFDQNRFVQ 60
Db 1 MAPTQTOAQREDGRPNSHRTLPERSGTVCRYKCNNSLPDTPFDPKPTYPFDQNRFVQ 60
Qy 61 YKATSLXKHQKHDLTPEPLGHTDLINDTYRIDPNVLDPDAEKLEELQAPTSKKR 120
Db 61 YKATSLXKHQKHDLTPEPLGHTDLINDTYRIDPNVLDPDAEKLEELQAPTSKKR 120
Qy 121 SQHAKVWPMRQBYISSEPNRYGISNEKPEVKIGVSQQTTEEYKRDQSQTAE 180
Db 121 SQHAKVWPMRQBYISSEPNRYGISNEKPEVKIGVSQQTTEEYKRDQSQTAE 180
Qy 181 KTFEDAKTSQSOHSKPRTPVEMPVFDKFQWINPQCAQVFDSDPAPDTSGAALEM 240
Db 181 KTFEDAKTSQSOHSKPRTPVEMPVFDKFQWINPQCAQVFDSDPAPDTSGAALEM 240
Qy 241 MSQAMIRGMDEBEEQFYFLPVVEETLKRKRDQEEMDYAAPDDYYDKTAREYNNWK 300
Db 241 MSQAMIRGMDEBEEQFYFLPVVEETLKRKRDQEEMDYAAPDDYYDKTAREYNNWK 300
Qy 301 NKASKGYENYFFREGGVYINLETRVLSKRAGAVQSGTNALLVVRDNEKE 360
Db 301 NKASKGYENYFFREGGVYINLETRVLSKRAGAVQSGTNALLVVRDNEKE 360
Qy 361 LEAPARKAQLENTPPEEEBEEMTEKEAGGSDEBOKGSSEKEGSSEDHGSSEER 420
Db 361 LEAPARKAQLENTPPEEEBEEMTEKEAGGSDEBOKGSSEKEGSSEDHGSSEER 420
Qy 421 EEGDRDEASDKSGREDESEDEARAARDKEIFGSDADSEDDSDPDEGQAOQGSDN 480
Db 421 EEGDRDEASDKSGREDESEDEARAARDKEIFGSDADSEDDSDPDEGQAOQGSDN 480
Qy 481 DSDGSNGGQSRSPHRSRASPSSPQGSEHSAQEDGSEAAADSSEADSDD 531
Db 481 DSDGSNGGQSRSPHRSRASPSSPQGSEHSAQEDGSEAAADSSEADSDD 531

RESULT 5
AAB93517 ID AAB93517 standard; protein; 531 AA.
XX AC AAB93517;
XX DT 26-JUN-2001 (first entry)
XX DE Human protein sequence SEQ ID NO:12853.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Query	Match	Score 99.3%; Best Local Similarity 99.4%; Matches 528; Conservative 0; Mismatches 3; Indels 0;	Score 2744; Pred. No. 3.9e-195; Mismatches 0; Indels 3; DR 0;
Qy	YKATSLXKHQKHDLTPEPLGHTDLINDTYRIDPNVLDPDAEKLEELQAPTSKKR 61	MAPTQTOAQREDGRPNSHRTLPERSGTVCRYKCNNSLPDTPFDPKPTYPFDQNRFVQ 60	
Db	SQHAKVWPMRQBYISSEPNRYGISNEKPEVKIGVSQQTTEEYKRDQSQTAE 61	MAPTQTOAQREDGRPNSHRTLPERSGTVCRYKCNNSLPDTPFDPKPTYPFDQNRFVQ 60	
Qy	1 MAPTQTOAQREDGRPNSHRTLPERSGTVCRYKCNNSLPDTPFDPKPTYPFDQNRFVQ 61	YKATSLXKHQKHDLTPEPLGHTDLINDTYRIDPNVLDPDAEKLEELQAPTSKKR 120	
Db	1 MAPTQTOAQREDGRPNSHRTLPERSGTVCRYKCNNSLPDTPFDPKPTYPFDQNRFVQ 61	SQHAKVWPMRQBTYISTEFNRYGISNEKPEVKIGVSQQTTEEYKRDQSQTAE 180	
Qy	121 KTFEDAKTSQSOHSKPRTPVEMPVFDKFQWINPQCAQVFDSDPAPDTSGAALEM 240	121 KTFEDAKTSQSOHSKPRTPVEMPVFDKFQWINPQCAQVFDSDPAPDTSGAALEM 240	
Db	181 KTFEDAKTSQSOHSKPRTPVEMPVFDKFQWINPQCAQVFDSDPAPDTSGAALEM 240	181 KTFEDAKTSQSOHSKPRTPVEMPVFDKFQWINPQCAQVFDSDPAPDTSGAALEM 240	

Qy	241	MSQMTRGMMDDEGQVAYFLPVETLKKRKDQEEMDYAPDDVYDYKIAREBYNNWYK	300
Db	241	MSQMTRGMMDDEGQVAYFLPVETLKKRKDQEEMDYAPDDVYDYKIAREBYNNWYK	300
Qy	301	NKASKGYENYFFPRREGGVYYNLETFVRLSKRAKGVQSTNALLVVKERDMNEKE	360
Db	301	NKASKGYENYFFPRREGGVYYNLETFVRLSKRAKGVQSTNALLVVKERDMNEKE	360
Qy	361	LEAQARYAQLENPEEEEEEETTEKEAGSDEEQPKGSSEKEGSEDEHSGSESER	420
Db	361	LEAQARYAQLENPEEEEEEETTEKEAGSDEEQPKGSSEKEGSEDEHSGSESER	420
Qy	421	EEDGRDEASDKSGGSEDESSDEAARAAARDKEEITFGSDADSEDDDRGQAQQGSDN	480
Db	421	EEDGRDEASDKSGGSEDESSDEAARAAARDKEEITFGSDADSEDDDRGQAQQGSDN	480
Qy	481	DSPGSNGGGQRSSHSRSASPFPGSRSQEDSEAAADSSEADSD	531
Db	481	DSPGSNGGGQRSSHSRSASPFPGSRSQEDSEAAADSSEADSD	531
RESULT 6			
	ABG19682	ABG19682 standard; protein; 553 AA.	
	ID	ABG19682	
	XX	ABG19682;	
	AC		
	XX		
	DT	18-FEB-2002 (first entry)	
	XX		
	DB	Novel human diagnostic protein #19673.	
	XX		
	KW	Human; chromosome mapping; gene therapy; forensic;	
	KW	food supplement; medical imaging; diagnostic; genetic disorder.	
	XX		
	OS	Homo sapiens.	
	XX		
	PN	WO200175067-A2.	
	XX		
	PD	11-OCT-2001.	
	XX		
	PF	2001IWO-US008631.	
	XX		
	PR	31-MAR-2000; 2000US-00540217.	
	PR	23-AUG-2000; 2000US-00649167.	
	XX		
	PA	(HYSE-) HYSEQ INC.	
	XX		
	PI	Drmanac RT, Liu C, Tang YT;	
	XX		
	WPI:	2001-639362/73.	
	DR	N-PSDB; AAS83869.	
	XX		
	PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.	
	XX		
	PS	Claim 20; SEQ ID NO 50041; 103pp; English.	
	XX		
CC	CC	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polynucleotide and polypeptide sequences have applications in	

Sequence 553 AA:						
	Query Match Best Local Similarity Matches	Match Conservative 6;	Score 96.2†; Prod. No. 9.2e-189;	DB 4; Length 553;	DB 1; Length 553;	SQ
Qy	1 MAPIIQOAOREDGHPRNSHRTLPERSGVCRYCNCNSLDPIDPFPKFIYTFDQNRFVQ	12 MAPIIQOAOREDGHPRNSHRTLPERSGVCRYCNCNSLDPIDPFPKFIYTFDQNRFVQ	60		60	
Db						
Qy	61 YKATSLKEQHKHDLLTEPDLGVTIDLNPDTYRIDPNVLDDPADEKLLEEIIQAPTSSSKR	72 YKATSLKEQHKHDLLTEPDLGVTIDLNPDTYRIDPNVLDDPADEKLLEEIIQAPTSSSKR	120		120	
Db						
Qy	121 SQQHAKVPPWNRKTEYISTEFNRYGINSNEKPEYIGVSYRQQTFTBEEIYKDRDSQTAAIE	132 SQQHAKVPPWNRKTEYISTEFNRYCIFHEKPEVKKGSTYQQFTTEEIIYKDRDSQTAAIE	180		191	
Db						
Qy	181 KTFPDAQKS-----ISOHYSKPRVTPVEVMMPVFDFKMMWINPCAQVTFDSDPAP	192 KTFPDAQKSVISVGGLGEARISQHYSKPRVTPVEVMMPVFDFKMMWINPCAQVTFDSDPAP	229		251	
Db						
Qy	230 KDTSGAALEMNSQAMRGMMDEBEGNQTVAYPLPEETIKKRKRQEEMDYAPPDVYDY	252 KDTSGAALEMNSQAMRGMMDEBEGNQTVAYPLPEETIKKRKRQEEMDYAPPDVYDY	289		311	
Db						
Qy	290 KIARETWNVNKRKASKGYBENTFFI FREQDGVTYNELETRVLRLSRRAKAGYQSGINALL	312 KIARETWNVNKRKASKGYBENYFPFIREQDGVTYNELETRVLRLSRRAKAGYQSGINALL	349		371	
Db						
Qy	350 VVKHRDNNEKELEAQEARKAOLENHEPEEEEEEMETEKEKEAGGSDDEEQEKSSSEKEGS	372 VVKHRDNNEKELEAQETRAQLENHEPEEEEEEMETEKEKEAGGSDDEEQEKSSSEKEGS	409		431	
Db						
Qy	410 EDEHSSESEREGDDEAASDGSGEDESSEDARAADEKEIIGSDAISEDDADSDDE	432 EDEHSSESEREGDDEAASDGSQDSSDYZKARAARDEKEIIGSDAISEDDADSDDE	469		491	
Db						
Qy	470 DRGQAQGSDNDSDGSNGGGQRSHRSASPPSGSEHSAQEGSEAAAASDSEADSD	492 DRGQAQGSDNDSDGSNGGGQRSHRSASPPSGSEHSAQENGSEAAAASDSEADSD	529		551	
Db						
Qy	530 SD 531	552 SD 553				
Db						
RESULT 7						
AAB56316	ID AAB56316	standard; protein; 473 AA.				
AC	AAB56316;					
XX	XX					
DB	Human secreted protein sequence encoded by gene 106 SEQ ID NO:410.					
XX						
KW	Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;					
KW	antiinflammatory; antiproliferative; cyrostatic; cardiotonic; vasoconstrictive;					
KW	cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;					
KW	fungicide; ophthalmological; gene therapy; pathological condition;					
KW	autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;					
KW	neoplasia; cardiovascular disorder; cardiac arrest; cerebral ischaemia;					
KW	cerebrovascular disorder; angiogenesis; nervous system disorder;					

KW	Alzheimer's disease; infection; ocular disorder; corneal infection;	Db	241 MSQAMIRGMMDDEBENQFVAYFLPVETLKRRKDQEEEMDYAPDDVYDKIAREYNWVK 300
XX	wound healing; skin aging; food additive; preservative.	Qy	301 NKASKGTYENTPFIFRGDGYYNELETRVLSKRAKAGYOSGNALLWKRDRNEKE 360
XX	Homo sapiens.	Db	301 NKASKGTYENTPFIFRGDGYYNELETRVLSKRAKAGYOSGNALLWKRDRNEKE 360
XX	WO2000070042-A1.	Qy	361 LPAQEARKAQLENHEPEEEEMETEEKEAGGSDEEQKSSEKEGSSEDEHSGESER 420
XX	23 - NOV - 2000.	Db	361 LPAQEARKAQLENHEPEEEEMETEEKEAGGSDEEQKSSEKEGSSEDEHSGESER 420
XX	11 - MAY - 2000; 2000000-US012788.	Qy	421 EFGDRDDEASDKSGEDDESEARARDEBIFGGSDADSDDEDRCQ 473
XX	13 - MAY - 1999; 99US-0134068P.	Db	421 EFGDRDDEASDKSGEDDESEARARDEBIFGGSDADSDDEDRCQ 473
(HUMA -) HUMAN GENOME SCI INC.	PA	PI	RESULT 8 ABB59163
XX	Rosen CA, Ruben SR, Moore PA, Young PE, Komatsoulis GA, Birse CE;	DB	ID ABB59163 standard; protein; 538 AA.
PI	Duan RD, Florence KA, Soppet DR,	XX	XX
XX	WPT; 2000-679828/66.	PT	AC ABB59163;
XX	Isolated nucleic acid molecule encoding a human secreted protein is used	XX	XX
PT	in preventing, treating or ameliorating a medical condition.	XX	DT 26-MAR-2002 (First entry)
XX	Disclosure; Page 1041-1042; 1065PP; English.	PS	CC Drosophila melanogaster polypeptide SEQ ID NO 4281.
XX	The Polynucleotide sequences given in AAC99818 to AAC9977 encode the	CC	DE Drosophila melanogaster; cell signalling; insecticide;
CC	human secreted proteins given in AAB5077 to AAB56362. Human secreted	CC	KW Drosophila; developmental biology; cell signalling; insecticide;
CC	proteins have activities based on the tissues and cells the genes are	CC	KW pharmaceutical.
CC	expressed in. Examples of activities include: immunosuppressive;	CC	XX Drosophila melanogaster.
CC	antiarthritic; antirheumatic; antiproliferative; cyrostatic; cardiant;	CC	OS PN WO200171042-A2.
CC	vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;	CC	XX 27-SEP-2001.
CC	virucide; fungicide; and ophthalmological. The human secreted	CC	XX 23-MAR-2001; 2001WO-US009231.
CC	polynucleotides and proteins can be used to prevent, treat or ameliorate	CC	PR 23-MAR-2000; 2000US-0191537P.
CC	a medical condition in e.g. humans, mice, rabbits, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological	CC	PR 11-JUL-2000; 2000US-00614150.
CC	condition or susceptibility to a pathological condition. Disorders which	CC	XX (PEKE) PB CORP NY.
CC	are diagnosed or treated include autoimmune diseases e.g. rheumatoid	CC	XX PI Venter JC, Adams M, Li PWD, Myers EW;
CC	arthritis, hyperproliferative disorders e.g. neoplasms of the breast or	CC	XX DR WPI; 2001-656860/75.
CC	liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular	CC	XX DR N-PSDB, AB130256.
CC	disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders	CC	XX New isolated nucleic acid detection reagent for detecting 1000 or more
CC	e.g. Alzheimer's disease, infections caused by bacteria, viruses and	CC	PT genes from Drosophila and for elucidating cell signalling and cell-cell
CC	fungi and ocular disorders e.g. corneal infection. The proteins can also	CC	PT interactions.
CC	be used to aid wound healing and epithelial cell proliferation, to	CC	XX SQ Disclosure; SEQ ID NO 4281; 21PP + Sequence Listing; English.
CC	prevent skin aging due to sunburn, to maintain organs before	CC	XX PS
CC	transplantation, for supporting cell culture of primary tissues, to	CC	CC The invention relates to an isolated nucleic acid detection reagent
CC	regenerate tissues and in chemotaxis. The proteins can also be used as a	CC	CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC	food additive or preservative to increase or decrease storage	CC	CC useful in developmental biology and in elucidating cell signalling and
CC	capabilities. AAC99809 to AAC99817 and AAC99809 represent sequences used	CC	CC cell-cell interactions in higher eukaryotes for the development of
CC	in the exemplification of the present invention	CC	CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX	Sequence 473 AA;	XX	CC discloses genomic DNA sequences (AB116176-AB13051), expressed DNA
Qy	89 1%; Score 2464; DB 3; Length 473;	Db	CC sequences (ABL01840-ABL1675) and the encoded proteins (ABBS737-CC ABB72072). The sequence data for this patent did not form part of the
Best Local Similarity 90 8%; Pred. No. 2..le-174;	Mismatches 0; Indels 0; Gaps 0;	Qy	CC printed specification, but was obtained in electronic format directly
Matches 472; Conservative 0; Mismatches 1;	Qy	CC from WIPO at ftp.wipo.int/pub/published_pct_sequences	
Db	1 MAPTIQTOAQREDGRPNRSHRTLERSGYCVCYCNSPDIPFDPKFTYPFDQNFVQ 60	XX SQ Sequence 538 AA;	
Db	1 MAPTIQTOAQREDGRPNRSHRTLERSGYCVCYCNSPDIPFDPKFTYPFDQNFVQ 60	Query Match 45.0%; Score 1244.5; DB 4; Length 538;	
Qy	61 YKATSLKQKHDLITEPDOLGTIDLNPDYRIDPNVLLDADEKLUFEEQAPTSSKR 120	Best Local Similarity 50.0%; Pred. No. 9.3e-34;	
Db	61 YKATSLKQKHDLITEPDOLGTIDLNPDYRIDPNVLLDADEKLUFEEQAPTSSKR 120	Matches 271; Conservative 66; Mismatches 172; Indel. 33; Caps 11;	
Qy	122. SQQHAKVWPMRKCYTISTEPNRYGINSNKEPEVKIGVSQKQPTEEETYKRDSDQTAIE 180	Qy 1 MAPTIQTOAQREDGRPNRSHRTLERSGYCVCYCNSPDIPFDPKFTYPFDQNFVQ 60	
Db	121. SQQHAKVWPMRKCYTISTEPNRYGINSNKEPEVKIGVSQKQPTEEETYKRDSDQTAIE 180	Db 1 NEPTINNSAVNSAAEKR-RPQROTTERSSEICRVKYCNPLPDPFDKFLQDPSHRSRFVQ 59	
Qy	181 KTFEDAKSISOHISKPRVTPVETEYKRDSDQTAIE 240	Qy 1 NEPTINNSAVNSAAEKR-RPQROTTERSSEICRVKYCNPLPDPFDKFLQDPSHRSRFVQ 59	
Db	181 KTFEDAKSISOHISKPRVTPVETEYKRDSDQTAIE 240		
Qy	241 MSQAMIRGMMDDEBENQFVAYFLPVETLKRRKDQEEEMDYAPDDVYDKIAREYNWVK 300		

Qy 61 YKATSLKEQKHDKLIDEPDGVTIDLINPDTYRIDPNVLDPADEKLLEBEIQAPTSSKR 120
 Db 60 YNPTSLERNFYDMLTEHDLGTVNDLRELYQADSMTLLDADEKLLEBEITLPDTSVR 119
 Qy 121 SQQHAKVPPWKRKTYISTENRNYGLSN-EKPEVKIGVSKVQQTBEETYKDRDSQTTAI 179
 Db 120 SRQHSRTVSNLRSKYSISTBOTRFQPQNLENIEAKVGYNTKSLREETLYLDREAQKAI 179
 Qy 180 EKTFEDQKSQPSKSPRTTPVEWMPVFDFKWMINPCQVTFPSDPAPKDTSGAAAL 239
 Db 180 EKTFSDTKSSELTGHTSKPNVTPVFLPIFFDTNWKFPCAVQFSDPAPAKNPVAQE 239
 Qy 240 MMSQAMIRGMNDEECNQFVAYFLPVETTLKCRKDQEEMDYAPDDVYDYLKARBYNNVV 299
 Db 240 EMSQAMIRGMNDEECNQFVAYFLPVETTLKCRKDQEEMDYAPDDVYDYLKARBYNNVV 299
 Qy 300 KTKASKGYENYFFIREGGCVYNELETIVRLSKRAGKGVQSGTNALLVVKHRDNNEK 359
 Db 300 KTKASKGYENYFFYMRQ-DGIVYNELETIVRLNKKRVKG-QQPNNTKLVVVKHRPLDSM 357
 Qy 360 ELEAQFARKAQLENLEPEE-----EEBON-----ETEB-----KEAGGD----- 395
 Db 358 EHMQBYREROLEVPGESEEELVETREEQMIGIETKETSEDAAVGROAAGSDPAQV 417
 Qy 396 --EEOEQKGSSEKECSSEDEHSGSSESEREGDRDEASDKSGSGEDESSEDDEAARDKEBI 453
 Db 418 ARDROQRSRSTRSES-SSSSGSGSRAASRSKSGRSRSGSRSRSGS- 475
 Qy 454 FGSDADSEDDADSDDEDRGQAQGGSDNSDSSGS-NGGGQBSRSRSRSASPPGSSERSAQ 512
 Db 476 SRSVSRSSRS 535
 Qy 513 ED 514
 Db 536 SD 537
 Qy RESULT⁹
 ID ABG19681 standard; protein; 133 AA.
 AC ABG19681;
 DT 18-FEB-2002 (First entry)
 XX Novel human diagnostic protein #19672.
 XX Human; chromosome mapping; gene mapping; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US008631.
 PR 31-MAR-2000; 2000US00540217.
 PR 23-AUG-2000; 2000US00619167.
 XX (HYSE-) HYSEQ INC.
 PA DRmanac RT, Liu C, Tang YT;
 PI XX
 DR N-PSDB; RAS83868.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 50040; 103pp; English.
 PS CC The invention relates to isolated Polynucleotide (I) and Polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, is molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG31377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 133 AA;

Qy Query Match 22.5%; Score 622; DB 4; Length 133;
 Best Local Similarity 64.4%; Pred. No. 2.6e-38;
 Matches 130; Conservative 0; Mismatches 2; Indels 70; Gaps 1;

Qy 273 RDQBEEMDYAPDYYDYLKIAREBNWNVYRNKA SKG YEN YFFIFREGDGVYMELETIVRL 332
 Db 1 RDQBEEMDYAPDYYDYLKIAREBNWNVYRNKA SKG YEN YFFIFREGDGVYMELETIVRL 332
 Qy 333 SKRRAKGVQSTGNALLVVKHRDMNEKELEAQEARKAQLENHEPEEEETEEKERAG 392
 Db 36 -----
 Qy 393 GSDEEQKGSSEKECSSEDEHSGSSESEREGDRDEASDKSGSGEDESSEDDEAARDKEB 452
 Qy 51 GSYFBEQBGSSSKKEGSEDEHSGSSESEREGDRDEASDKSGSGEDESSEDDEAARDKEB 110
 Db 453 IFGSDADSEDDADSDDEDRGQA 474
 Db 111 IFGSDADSEDDADSDDEDRGQA 132

RESULT 10
 AAQ03326 standard; protein; 115 AA.
 ID AAQ03326;
 AC AAQ03326;
 XX DT 06-OCT-2000 (first entry)
 XX DE Human secreted protein, SEQ ID NO: 7407.
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX OS Homo sapiens.
 XX PN EP103301-A2.
 XX PD 06-SEP-2000.
 XX PR 21-FEB-2000; 20000BP-00200610.
 XX PR 26-FEB-1999; 99US-0122481P.
 XX (GEST) GENSET.

PI	Dumas Milne Edwards J;	Ducleert A,	Giordano J;	PT responsible for genetic disorders or other traits and to assess biodiversity.
XX	WPI; 2000-500381/45.			XX Claim 20; SEQ ID NO 49771; 103pp; English.
DR	N-PSDB; AAC03332.			PS
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.			XX
XX	Claim 13; SEQ ID NO 7407; 71pp + Sequence Listing; English.			CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, Polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (III) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (III) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The peptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABS00010-AB30377 Represents novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
PS	The present sequence is a polypeptide encoded by one of a large number of cDNA sequences derived from mRNAs encoding secreted proteins. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNAs sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors			CC
XX	Sequence 115 AA;			CC
SQ	Query Match Score 21.5%; Best Local Similarity 96.5%; Pred. No. 2.1e-36; Matches 111; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			CC
Qy	1 MAPIQTQREDEHPRNPSHRTLPERSGIVCRVYCNNSLDIPFDPKPFTYPFDQNRFVQ 60			CC
Db	1 MAPIQTQREDEHPRNPSHRTLPXXSEVCRVYCNNSLDIPFDPKPFTYPFDQNRFVQ 60			CC
Qy	61 YKATSLXEKHKHDLTTEPLPGVTIDLINPDTYRIDPNVLLDADEKULEBIOAP 115			CC
Db	61 YKATSLXEKHKHDLTTEPLPGVTIDLINPDTYRIDPNVLLDADEKULEBIOAP 115			CC
RESULT 11				CC
ABG19412	ABG19412 standard protein; 475 AA.			CC
XX				XX
AC	ABG19412;			XX
XX				XX
DT	13-FEB-2002 (first entry)			XX
DE	Novel human diagnostic protein #19403.			XX
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.			XX
KW				XX
XX				XX
OS	Homo sapiens.			XX
XX	WO200175067-A2.			XX
PD	11-OCT-2001.			XX
XX	(HYSEQ INC.			XX
PF	30-MAR-2001; 2001WO-US008631.			XX
XX				XX
PR	31-MAR-2000; 2000US-00540217.			XX
PR	23-AUG-2000; 2000US-00549167.			XX
XX				XX
PA	(HYSEQ INC.			XX
XX				XX
PF	30-MAR-2003; 2003WO-US030720.			XX
XX				XX
DR	02-OCT-2002; 2002US-0416186P.			XX
XX				XX
PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations			XX

Page 9

ID	ABR53245	standard; protein; 445 AA.
XX		
AC	ABR53245;	
XX		
DT	20-JUN-2003	(first entry)
XX		
DE		Protein sequence #SEQ ID: 1335.
XX		
KW		Multiprotein complex; eukaryote; drug target; diagnosis

XX Saccharomyces cerevisiae.
OS XX EP1258494-A1.
PN XX PD 20-NOV-2002.
XX 20-DEC-2001; 2001EP-00130253.
PF XX

PR 15-MAY-2001; 2001EP-001117/4.
XX (CELL-) CELLZONE AG.
PA XX
PA XX
PI XX

Disclosure: SRO ID NO 1355: 17pp + Sequence Listing: Enc
 PI Marioch M, Schultz JN, Superti-Furga GD;
 XX
 DR WPI: 2003-250078/25.
 N-PSDB: ACC61287.
 XX
 PT New isolated protein complexes useful for diagnosing a c
 disorder, or as a target for an active agent of a pharm
 preferable a drug targeted in the treatment or prevention
 disorder.
 PT
 PT
 PT
 XX

The invention relates to multiprotein complexes from eukaryotes. Proteins of the invention and DNA sequences encoding them are given in records ABR52568-ABR53903 and ACCG610-ACCG6194 respectively. The complexes are obtainable by using a protein as a bait and isolating the set of proteins which is attached thereto from cells. Such protein complexes may comprise up to 30 distinct proteins. Protein complexes of the invention are useful for diagnosing a disease or disorder, or as a target for an active agent of a pharmaceutical, preferably a drug target in the treatment or prevention of a disease or disorder. Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office. The complete document is available on CD-ROM.

Sequence 445 AA;	
query Match	9.2t;
best Local Similarity	22.1t;
Pred. No.	3.2e-10;
Matches	112;
Conservative	96;
Mismatches	184;
Indels	114;
Gaps	21;
23 LPERSGVYVCURVYCNSQLDIPFDPKFITYP-----FDQNRFVQYKATSLEKQHK 71	:
1 MSKKQEVTAPIKYQNSLSPVQLQPKLVLVPESPETNAQSQLINSLYIKITNTNLQQ-- 58	:
72 HDLTLTEPDGLVTIDL-----NPDTYRILDPNVLLDPADEKLLLEEIDQAPTSKRS 121	:
59 ---DELGMPFDLMKPGPLMLKLDKLGYFD-NVKLDKDRILLRD---PRDRLT 108	:
122 CQHAKVYVPMRKNEYISTEFNRYQGINSKEPVKGIVSVKQOFTEBEEIYKDRDSQ---- 175	:
109 KIDISKTYFLRLRTBEVNTTAAHDNTSLKRKRRL-----DDGQSDDENLDV 154	:
176 ---ITAIEKTPEDAQKSISQHYSKPRVTPVEVMVMPDFQWINKINCAQVIFPDSDFPAPKDT 232	:
155 NHISRSVEGTENTKTDK--WQHPYTKGKVKWKKWNLDPD---TASMDQVTF----ILKF 203	:
223 SGAALEMMSQAMIR-GM---MDEBGNQPVAYFLEPVEETLKGRKRDQEEMDYAPPDVYD 288	:
204 MGSASLTDKEKSSLNTGFRPVEELDEDWISMYATDKDOSAILENLEKGMDEMDDDSHE 263	:
289 ---YKLAEBYNNWVTKASKGKVYENFFI FREGDGV-YYNELERTRVLSGRKAGVQSG 344	:
264 GLYKFKIRIDYMKQVAEKPWTE-LATRLNDKOGIATYKPLRSKTELRRRVRND1IKP- 321	:
345 TNALLVYKH-----RDMNEKELEAQBAEKQLEN----HPEEEEEEEMSTEKK 389	:
322 ---LVKEHDIDQNLNTLRNPSTKEANTRDLRNMKFDPINFATVEDDEDEBEPQPDVKK 377	:
390 EAGGSDEHQEKGSSESERGSSEDEHSGSESBREEGDORDEASDKSGGGEDESSEDEAARAD 449	:
378 ESEFG--DSKTEGSEQEGENERDEEIKQEKENEQ----DEENKDENEAADT 422	:
450 KERIFGSDADSEADSDDEDRGQAQ 475	:
423 PET---SDAVHTQKPEEEBKETQEE 445	:
ULT 14	
63670 ADK63670 standard; protein; 445 AA.	
Unidentified.	
06-MAY-2004 (first entry)	
Disease treating protein complex-derived protein #818.	

PD	27-AUG-2003.	Db	322 ---LYKEHDIDOLNVTLNRNPSTKEANIRDKLRMKFDINPATVDEDDEDEEQPEDVKK	377
XX	20-DEC-2002; 2002EP-00102902.	QY	390 EAGGSDEEQEKSSESSSEKGSEDEHSSESEREGDRDEASDKSGSGDESESDEAARARD	449
PP	XX	DB	378 ESEG---DSKTEGSEQRGENEDEBIIQEKENEQ-----DEENIQDENRAADT	422
PR	20-DEC-2001; 2001EP-0010253.	QY	450 KEPIFGSDADDDDDDEQDQAAQ	475
XX	(CELL-) CELIZOMB AG.	DB	423 PET---SDAVTEQKPEEKETLOEE	445
PA	Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J, Krause R, Merino A, Bauch A;			
PI	Michioch M, Grandi P, Krause U, Merino A, Bauch A;			
PI	Michon A, Leutwein C,			
XX	RICK J;			
DR	WPI; 2003-638460/61.			
N-PSDB;	ADK63671.			
XX	New proteins and protein complexes from eukaryotes, useful as targets in drug screening, or in diagnosing or screening for the presence of a disease or disorder, or a predisposition for developing a disease or disorder in a subject.			
XX	Disclosure; SEQ ID NO 1635; 13pp; English.			
XX	The invention relates to novel protein complexes comprising a first and a second protein, or its derivative, fragment, homologue or variant. The proteins are selected from given protein complexes, which are not defined in the specification. The variants are encoded by nucleic acids that hybridize to the nucleic acids encoding the proteins under low stringency conditions. The protein complexes are useful as targets for an active agent of a pharmaceutical. These protein complexes are particularly useful as drugs targets for the treatment or preventing of a disease or disorder. The complexes and methods above are useful in diagnosing or screening for the presence of a disease or disorder or a predisposition for developing a disease or disorder in a subject. These are also useful in screening for a drug for treatment or prevention of a disease or disorder. The molecule that modulates the amount, activity or protein components of the complex is useful for the manufacture of a medicament for the treatment or prevention of a disease or disorder. This sequence corresponds to a protein of the invention. (Note: the sequence data for this patent did not form part of the printed specification but was obtained from the EPO in electronic format).			
CC	Sequence 445 AA;			
CC	Query Match Score 250; DB 7; Length 445;			
CC	Best Local Similarity 22.1%; Pred. No. 3.2e-10;			
CC	Matches 112; Conservative 36; Mismatches 184; Indels 114; Gaps 21;			
QY	23 LPERSGVVRCKYKNSLPLDIPDPKPFITYP-----FDQNRFQYKATSLEKOKH	71		
DB	1 MSKKOBYIADIKYKNSLPLPQLPKJLVLVPESPTNADSQLNSLYKTNTNLIGC--	58		
QY	72 HDLITEPDGVTDLJ-----NPDTYRDPNVLDPADEKLIEFQIAPTSSKRS	121		
DB	59 ----DEDLGMPVDMKFPGLLNKLDSLKLYGFD-NVKLDKDRLR----PRIDRL	108		
QY	122 QHQAKVPPWRKRNLTISTPNRYGISNEPKEVKIGVSYQQFTEEBIYKRDQS-----	175		
DB	109 KTDISKYTFERTRTVEVNTIAAHNTSLPKRRL-----DGGSDDDENIV	154		
QY	176 --ITAIEKTFEDAKSISOHYSKPRVTYVVMVPFDPFKMWNPQAOYTFDSDPAPKTD	232		
DB	155 NHISRVEGTFNRTDK--WQHPVKGVKVNWKWDLLPD---TASMDQYF---ILKF	203		
QY	233 SGAAALEMMQSOAMIR-GM---MDEEGNOVAYAFVLPVEETLUKKRQDEEMDYAPDDVYD	288		
DB	204 MGSA8LDTGEKSKSNTGIPRPVLEEDEVTSYMATDHKOSAILENELLGMDMDDDSHE	263		
QY	289 --YKIAARETWNWNKAKSGYBENYFFFREGDGV-YTNELETRVLSKRRAKAGVQS	344		
DB	264 GKTYKEFRKRTYDQKVAPEMTS-LATIINDKGIATYKPLRSKRRYNDLKPK-	321		
QY	345 TNALLVVKH-----RDMNEKELEAQEPARKQLEN-----HEPPEEEBEMETEK	389		

CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp://wipo.int/pub/published_pct_sequences

xx Sequence 1633 AA;

Query	Match	Score	Length
Qy	Best Local Similarity	8.6%	1633;
Matches	21	21	2.5e-08;
Db	Local Similarity	8.5%	238;
Qy	Conservative	85;	Mismatches 238;
Db	Matches	21	Indels 165;
Qy	Score	237.5;	Gaps 25;
48	DB	DB	DB
YVTLKDSNNRELQRVTDQSGHYZQFDNLQNGT--YTFEPATPDPNTTPSPANNSTNDAIDS	707	707	707
ADBKLLBEEIQAQPTSSRSQHAKV-----VPWWRKRTETYSTEFRYGI--SNEKPEVK	103	103	103
DGERDGTRKVVVAKGTINNAADMIMVTDTGFIYLTPKYNGDYTWEDTNKGQDDNEKGISN	765	765	765
IGVSVKQ-----QFTEPEIY-----KDRDSQ	155	155	155
VKVTLKKNKGNDTIGTTDNGSKYBFTGLENDYTFEFTEPGYPTKQNSGSDGKDSN	825	825	825
ITAIETKFEDAQKSISOHYSKSPRVTPEVIMPVFDKMMWINPCAOVIFPDS-----	176	176	176
GTKTTTVTKDADNKTTDGSFYKPKIYN-----LGDY-WMDETNKGDIODSEKGTSGYK	885	885	885
-APKDTSGAA-----ALEMMQSAMIRGMDEBENQFAYFIVEETLKKRKRDQEEMNDY-	228	228	228
VTLKDKNQNAIGTTTIDASHYQPKGL--ENGSYTVEPETSGYPTKANSQDITVDSN	937	937	937
-APDDVYD--YKJAR---EYNNWNVNQK-----ASKGY-----307	282	282	282
GTTTGTLINGADNLTIDSGFYKTPKISVGDTWEDTNKGQDDNEKGISVYKTLKDEK	995	995	995
-EENYFPFREGD-GVYVNELETRVRLSKRAGVQSGTNALLVVKHRDMN	308	308	308
GNI1STTTDENGKYQFDNLDSGNYIHFKEPEGMQTTRANS-----NDD	1055	1055	1055
EKELEAQEARKAQDNH-----EPPEEEBFEMETEBKAGSD	358	358	358
EKDADGEDVR-VT1DHDFSDYDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD	1101	1101	1101
EEQEKGSSSEKEGSSEHEGSESEERE-EGRDRDEASDK-SGSGEDESSEDARAARDKEE	396	396	396
SDDSD	1160	1160	1160
FGSDADSEDDADSDDDERGQAQGGSDNDSDSGSGNQQQRSHRSASPPFGSGSEHSQAQ	454	454	454
SDDSDADSDADSDADSDADSDADSDADSDADSDADSDADSDADSDADSDADSDADSD	1220	1220	1220
DGSEAAASDS-SEADSDSD	514	514	514
DADSDSDSDSDSDADSDSD	1277	1277	1277

Search completed: September 26, 2005, 13:51:43
 Job time : 171 secs

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2764	100.0	531	2	Q9H166		O9h166 homo sapien
2	2744	99.3	531	2	Q9NUU9		O9nuu9 homo sapien
3	2718	98.3	535	2	Q8K2T8		Q8k2t8 mus musculu
4	2708	98.0	535	2	Q9JU99		Q9ju99 mus musculu
5	2531	91.6	510	2	Q75239		Q75239 homo sapien
6	2183	79.0	520	2	Q8P2Y1		Q8p2y1 xenopus tro
7	1995	72.2	485	2	Q8NTH5		Q8nth5 homo sapien
8	1984	71.8	377	2	Q9CS63		Q9cs63 mus musculu
9	1935	70.0	407	2	Q8F851		Q6851 xenopus lae
10	1244	45.0	538	2	Q9VN55		Q9vn55 drosophila
11	1129	40.8	468	2	Q8PKX4		Q7pxz3 anophales g
12	645	23.3	425	2	P90783		P90783 caenorhabdi
13	372	13.5	589	2	Q8RW91		Q8rw91 arabidopsis
14	361	13.1	593	2	Q9MMA4		Q9maa4 arabidopsis
15	358	13.0	451	2	Q8DZ92		Q6zg32 oryza sativ
16	335	12.1	547	2	Q9CA82		Q9ca82 arabidopsis
17	286	10.4	386	2	Q6C509		Q6c509 yarrowia li
18	280	10.1	457	2	Q9US06		Q9usb6 schizosacch
19	265	9.6	791	2	Q9DGJ1		Q9dgj1 fugu rubrip
20	257	9.3	571	2	Q8MTN7		Q8mbn7 trichinella
21	253	9.2	445	1	PAP1_YEAST		P3851 saccharomy
22	251	9.1	538	2	Q9BT15		Q9et15 mus musculu
23	250	9.1	458	2	Q6BRT93		Q6br93 debaromyces
24	245	8.9	784	2	Q7LZ90		Q71293 torpedo cal
25	245	8.9	1848	2	Q7RGP8		Q7rgp8 plasmodium
26	244	8.8	934	2	Q9GM32		Q9gm32 bos taurus
27	243	8.8	438	2	Q6FXJ9		Q6fxj9 candida gla
28	241	8.7	792	2	Q9YTL7		Q9yl7 ateline her
29	240	8.7	1394	1	Q9NG4_BOVIN		Q28181 bos taurus
30	238	8.6	946	2	Q7QC53		Q7qg33 anopheles g
31	238	8.6	613	2	Q6UDM5		Q6udm5 brachydanio

DT	01-OCT-2002 (T=EMBLref.)	22; Last sequence update)	Db	361 LEAQARKAQLENHEPEEEEMEAEBKEAGGSDEEQEKSSEKEGSEDEHSGSESDR 420
DT	01-MAR-2004 (T=EMBLref.)	26; Last annotation update)	Qy	421 EEGDRDEASDKSGGEDESSEDEARAARDKEIFGSDADSEDDADSDDBDRCQAGGSDN 480
DE	RIKEN cDNA 5730511K23.		Db	421 SEGDRDEASDKSGGEDESSEDEARAARDKEIFGSDADSEDDADSDDBDRCQAHGSDN 480
GN	Name=s730511K23.		Qy	481 DSDGSGNGGQR ---SRSHSRSASPFPGSEHSAQEDGSEAAASSEADSDSD 531
OS	Mus musculus (mouse).		Db	481 DSDGSGNGGQR ---SRSHSRSASPFPGSEHSAQEDGSEAAASSEADSDSD 531
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Qy	DSDGSGNGGQR ---SRSHSRSASPFPGSEHSAQEDGSEAAASSEADSDSD 531
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Db	DSDGSGNGGQR ---SRSHSRSASPFPGSEHSAQEDGSEAAASSEADSDSD 531
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CZECH II; TISSUE=Mammary tumor;			
RC	MEDLINE=23388257; PubMed=124771932; DOI=10.1073/pnas.2412603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	RESULT 4	Q9JU99	PRELIMINARY;
RA	Klauserer A.D., Collison F.S., Wagner L., Schuler G.D.,		Q9JU99	PRT;
RA	Altschul S.F., Zeeberg B., Bauerow K.H., Schaefer C.F.,		AC	535 AA.
RA	Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		ID	
RA	Diatchenko L., Marszina K., Farmer A.A., Rubin G.M., Hong L.,		Q9JU99;	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Cabavant T.L., Scheetz T.E.,		AC	
RA	Brownstein M.J., Usdin T.B., Yoshiyuki S., Carninci P., Prange C.,		DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
RA	Raha S.S., Loquebadou N.R., Peters G.J., Abramson R.D., Mullahay S.J.,		DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		DB	Mus musculus brain cDNA, clone MNcb-6444, similar to Homo sapiens cDNA
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		DE	FLJ11123, clone PLACE1006167.
RA	Fahey J., Heilton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		DE	Name=5730511K23rik;
RA	Whalesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		GN	OS
RA	Rodriguez A.C., Grinwood J.W., Touchman J.W., Green E.D., Dickson M.C.,		GN	Os musculus (Mouse).
RA	Krzewinski M.I., Skalska U., Smajlau D.E., Schein J.E.,		GN	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
RA	Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human		GN	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae
RT	and mouse cDNA sequences";		GN	Mus.
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		GN	[1]
RL	[2]		RN	SEQUENCE FROM N.A.
RN	RP		RC	SEQUENCE FROM N.A.
RC	STRAIN=CZECH II; TISSUE=Mammary tumor;		Osada N., Kubuda J., Tanuma R., Ito A., Hirata M., Sugano S., Hashimoto K.,	
RA	Strausberg R.; Submitted: 05-2002 to the ENBL/GenBank/DBJ databases.		RA	RA
DR	EMBL; BC029843; AAH9843_1; -;		RA	Submitted (APR-2000) to the ENBL/GenBank/DBJ databases.
DR	InterPro: IPR007133; Pfam: 1;		DR	EMBL; AB041615; BAA95058_1; -;
DR	PFAM: PF03985; Pfam: 1;		DR	MGD; MGJ:1923988; 5730511K23rik.
DR	InterPro: IPR007133; Pfam: 1;		DR	InterPro: IPR007133; Pfam: 1;
DR	PFAM: PF03985; Pfaf1; 1;		DR	PFAM: PF03985; Pfaf1; 1;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro: IPR007133; Pfam: 1;		DR	SEQUENCE 535 AA;
DR	PFAM: PF03985; Pfaf1; 1;		DR	SEQUENCE 535 AA;
DR	SEQUENCE 535 AA;		DR	SEQUENCE 535 AA;
DR	MGD; MGJ:1923988; 5730511K23rik.		DR	SEQUENCE 535 AA;
DR	InterPro:			

Qy	481 DSDGSNGGQR----SRSRSASPFPGSEHSAEQEDSEAAADSSEADSD 531	529 DSD 531
Db	481 DSDSSDGQSERSQRSRSRSASPFPGSEHSAEQEDSEAAADSSEADSD 535	508 DSD 510
RESULT 5		
ID	075239; PRELIMINARY;	PRT; 510 AA.
AC	075239;	PRT; PRELIMINARY;
DT	01-NOV-1998 (TREMBrel. 08, Created)	DT 05-JUL-2004 (TREMBrel. 27, Last sequence update)
DT	01-JUN-2003 (TREMBrel. 24, Last annotation update)	DT 05-JUL-2004 (TREMBrel. 27, Last annotation update)
DB	F2149_1.	DE Hypothetical protein MGCG6249.
OS	Homo sapiens (Human).	GN Name=MGCG6249;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
NCBI_TaxID	96006; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OX	[1]	OC Xenopidae; Xenopus.
RN	NCBI_TaxID=8364;	OX NCBI_TaxID=8364;
RP	SEQUENCE FROM N.A.	RN [1] SEQUENCE FROM N.A.
RA	Lamerdin J.E., McReady P.M., Skowronski E., Adamson A.W., Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S., Phan H., Velasco N., Ho L., Regala W., Barnes J., Dangaran L., Poundstone P., Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andrade T., Frankheim M., Amico-Keller G.D., Kronmiller B., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kobayashi A., Olsen A.S., Carrano A.V.; Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.	RP TISSUE=Embryo; RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	EMBL; Acc005239; AAC25503.1. -.	RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Heieh P., RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarante P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., DR Fahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schwartz J., Myers R.M., Butterfield Y.S., RA Krzywinski M.I., Shalska U., Smailius D.B., Schnerch A., Schein J.E., RA Jones S.J., Marrs M.A.; Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [2]
SQ	SEQUENCE 510 AA; 57.66 MW; CACE73EDC7290CE8 CRC64;	RN SEQUENCE FROM N.A.
Qy	91.6%; Score: 2531.5; DB 2; Length: 510; Best Local Similarity: 91.7%; Pred. No. 4..4-112; Matches 498; Conservative 0; Indels 45; Gaps 3;	RP TISSUE=Embryo; RC Klein S., Gerhard D.S.; Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
Db	1 MAPIQTQQAQREDHEPNSHRTLPERSGVCRYKCNLSLPDIPDPKPTYFDQNRFVQ 60	RL EMBL; BC054253; AAH64253.1. -.
Qy	61 YKATSLEKQHKHDLTEPGLVTIDLNPDTYRIDPNVLDPDAEKLEELQAPTSKKR 120	DR InterPro; IPR007133; Pati.
Db	51 YKATSLEKQHKHDLTEPGLVTIDLNPDTYRIDPNVLDPDAEKLEELQAPTSKKR 110	DR Pfam; PF03985; Patf1. 1.
Qy	121 SQQAKVWMPRKELEYISTEFNGISNEKPEPKIGSVKQFTEEIYKORDSQTAIE 180	KW Hypothetical protein.
Db	111 SQQAKVWMPRKELEYISTEFNGISNEKPEPKIGSVKQFTEEIYKORDSQTAIE 170	SO SEQUENCE 520 AA; 59054 MW; 76D526C23C459A7 CRC64;
Qy	181 KTFEDAOKS1ISONSKPRTVPEVMVPDFKRNINCPAQFDSDBAPDTSQAALEM 240	Query Match 79.0%; Score: 2183.5; DB 2; Length: 520; Best Local Similarity 80.2%; Pred. No. 1..2e-95; Matches 429; Conservative 40; Mismatches 45; Indels 21; Gaps 8;
Db	171 KTFEDAQS---MWINPCQVFDPAFDTSGAALM 207	1 MAPIQTQQAQREDHEPNSHRTLPERSGVCRYKCNLSLPDIPDPKPTYFDQNRFVQ 60
Qy	241 MSQMIRGMDEEQQFYAYFLPVVEETLRKRDQEEMDYAAPDDYDCKAREYWNVK 300	1 MAPIQTQQAQREDHEPNSHRTLPERSGVCRYKCNLSLPDIPDPKPTYFDQNRFVQ 60
Db	208 MSQMIRGMDEEQQFYAYFLPVVEETLRKRDQEEMDYAAPDDYDCKAREYWNVK 267	1 MAPIQTQQAQREDHEPNSHRTLPERSGVCRYKCNLSLPDIPDPKPTYFDQNRFVQ 60
Qy	301 NKASKGKYYFETPREQGIVYNNELLR----VLSKSRKAGVQSGTNAL 348	61 YKATSLIEKOKHDLTBPDIGVTIDLNPDTYRIDPNVLDPDAEKLEELQAPTSKKR 120
Db	268 NKASKGKYYFETPREQGIVYNNELLR----VLSKSRKAGVQSGTNAL 327	61 YKATSLIEKOKHDLTBPDIGVTIDLNPDTYRIDPNVLDPDAEKLEELQAPTSKKR 120
Qy	349 LIVYGRHDNMEKELAQBARAKQLENHEPEEEEMETEKAAGGSDBEQKGSSSEREG 408	121 SQQAKVWMPRKELEYISTEFNGISNEKPEPKIGSVKQFTEEIYKORDSQTAIE 180
Db	328 LIVYGRHDNMEKELAQBARAKQLENHEPEEEEMETEKAAGGSDBEQKGSSSEREG 387	121 SQQAKVWMPRKELEYISTEFNGISNEKPEPKIGSVKQFTEEIYKORDSQTAIE 180
Qy	409 SEDEHSGSSESEREGDRDBASDKSGSEDESSEDAARDKEIFGSDADEDDSD 468	181 KTFEDAOKS1ISONSKPRTVPEVMVPDFKRNINCPAQFDSDBAPDTSQAALEM 240
Db	388 SEDEHSGSSESEREGDRDBASDKSGDESESSEDAARDKEIFGSDADEDDSD 447	181 KTFEDAOKS1ISONSKPRTVPEVMVPDFKRNINCPAQFDSDBAPDTSQAALEM 240
Qy	469 EDRCQAQCSNDNSDGSNGGGORSRSRSASPPSGSEHSAQEDSEAAADSSEADS 528	Db
Db	448 EDRCQAQCSNDNSDGSNGGGORSRSRSASPPSGSEHSAQEDSEAAADSSEADS 507	

241	MSQAMIRGMMDREGNQFTAYFLPVEETIKKRKDQEEMDYAPDDVYDVKAREYNWVK	300	QY	301 NKASKGYYENPFIREFGVYNTLETRVLSSKRAKGVOSETNALLVVKHRDNNEKE	360
241	MSQAMIRGMMDREGNQFTAYFLPGEETIKKRKDQEEMDYAPDDVYDVKAREYNWVK	300	Db	291 NKASKGYYBNPFIREFGVYNTLETRVLSSKRAKGVOSETNALLVVKHRDNNEKE	350
301	NKASKGYYENPFIREFGVYNELETRVLSSKRAKGVOSETNALLVVKHRDNNEKE	360	QY	361 LEQEARQAQLENHEPEEESEEMTEKEAGGS	394
301	NKASKGYYENPFIREFGVYNELETRVLSSKRAKGVOSETNALLVVKHRDNNEKE	360	Db	351 LEQEARQAQLENHEPEEESEEMTEKEAGGS	384
RESULT 8					
361	LEQEARQAQLENHEPEEESEEMTEKEAGGS-EEQEKGSSEKEGSDEDEHSGSSESE	419	Q9CS63	PRELIMINARY;	PRT;
361	LEQEARQAQLENHEPEEESEEMTEKEAGGSDEBEGSEBQSGEBSB	417	ID Q9CS63	AC 09CS63;	
420	REEGDRDEASDKSGSGEDESSEDEARAARDKEEFIGSDADSDDEDDRGQAGCSD	479	DT 01-JUN-2001	(TREMBLrel.	17, Created)
418	REEAEPEEKEDE-BEKESSSEEDRAARDKEEFIGSD---DDDSDED-GPNESQD	466	DT 01-JUN-2003	(TREMBLrel.	17, Last sequence update)
480	NDSGSGNG- -GGQRSPRSRSASAPPFPSGEHSAQBQDGESEAADSDDSD	531	DT 01-JUN-2003	(TREMBLrel.	24, Last annotation update)
467	GB-DSGSDDEEBKGQGRRSRSASSSPF-QSDHSQQENEDQSASDQCGSSGTSQSD	518	DB Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone: 5730511K23 product: PD2 PROTEIN (HYPOTHETICAL 60.0 kDa DE PROTEIN) homolog (Fragment).		
SEQUENCE FROM N.A.					
[1]	Q8N7HS	PRELIMINARY;	RN [1]	SEQUENCE FROM N.A.	
C	Q8N7HS;	PRT: 485 AA.	RP	SEQUENCE FROM N.A.	
C	Q8N7HS; 01-OCT-2002 (TREMBLrel.	Created)	RC	STRAIN=C57BL/6J; TISSUE=Whole body;	
C	01-OCT-2002 (TREMBLrel.	22, Last sequence update)	RX	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;	
b	01-JUN-2003 (TREMBLrel.	24, Last annotation update)	RA	Carninci P., Hayashizaki Y.;	
b	01-JUN-2003 (TREMBLrel.	Hypothetical protein FLJ2557.	RT	"High-efficiency full-length cDNA cloning.";	
b	Homo sapiens (Human).		RL	Meth. Enzymol. 303:19-44 (1999).	
b	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		RN [2]	SEQUENCE FROM N.A.	
b	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		RC	STRAIN=C57BL/6J; TISSUE=Whole body;	
b	NCBI_TaxID=9606; [1]		RX	MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;	
b	SEQUENCE FROM N.A.		RA	RIKEN FANTOM Consortium;	
b	TISSUE=Thyroid;		RT	"Functional annotation of a full-length mouse cDNA collection.";	
b	Minomura K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,		RL	Nature 409:685-690 (2001).	
b	Furiya T., Takahashi M., Kikkawa B., Omura Y., Abe K., Kamihara K.,		RN [3]	SEQUENCE FROM N.A.	
b	Katsuta N., Sato K., Tanikawa M., Suzuki Y., Hata H.,		RP	STRAIN=C57BL/6J; TISSUE=Whole body;	
b	Nakagawa K., Mizuno S., Morinaga M., Yamazaki M., Sugiyama T.,		RC	The FANTOM Consortium,	
b	Irie R., Otsuki T., Sato H., Niishikawa T., Sugiyama A., Kawakami B.,		RA	the RIKEN Genome Exploration Research Group Phase I & II Team;	
b	Nagai K., Isogai T., Sugano S.,		RT	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";	
b	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.		RT	Nature 420:563-573 (2002).	
b	EMBL; AK094423; BAC05305.1; IntePro; IPR007133; Pfam1.		RL	Genome Res. 10:1617-1630 (2000).	
b	PFAM; PF03985; Paf1.		[4]	SEQUENCE FROM N.A.	
b	SEQUENCE 485 AA;	55501 MW;	RP	STRAIN=C57BL/6J; TISSUE=Whole body;	
b	5F4A1AC99142C1D CRC64;		RC	MEDLINE=11042159; PubMed=20499374; DOI=10.1101/gr.145100;	
b	Query Match 72.2%; Best Local Similarity 97.2%; Matches 383; Conservative	Score 1995; DB 2; Pred. No. 9..3e-87; Indels 0; Gaps 1;	RA	Carninci P., Shibata K., Sugahara Y., Itoh M., Itoh M.,	
b	1 MAPIQOQAQREDGHPRNSHRLPERSGVCRVKYCNLSLDIFDPKPXTIPFDQNRFVQ	60	RA	Yamamoto H., Akiyama J., Nishi K., Kitasuna T., Tashiro H., Harada S., Nagaoaka S., Sasaki N., Carninci P.,	
b	1 MAPTIQOQAQREGH-----RGGVCRVKYCNLSLDIFDPKPXTIPFDQNRFVQ	50	RA	Sumi N., Ishii Y., Nakamura S., Nishine T., Harada A., Saito R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,	
b	61 YKATSLBKQKHDLILTPDGLYTIDLNPDTRIDPNVLIDPADEKULEEIQAPTSKKR	120	RA	Fujiwaka S., Inoue K., Togawa Y., Iizawa M., Ohara E., Watanuki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;	
b	51 YKATSLBKQKHDLILTPDGLYTIDLNPDTRIDPNVLIDPADEKULEEIQAPTSKKR	110	RA	Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;	
b	121 SQHQAKTVPRKTEYSTENPRYGINKEPEVKIGYSVKQFTEIYKORDSQITAI	180	RA	RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";	
b	111 SQHQAKTVPRKTEYSTENPRYGINKEPEVKIGYSVKQFTEIYKORDSQITAI	170	RT	Genome Res. 10:1757-1771 (2000).	
b	181 KTFEDAKSISQHYSKRPVTPYVMPYFPDFKWINCPAQVTFDSDPAPKOTSGAALEM	240	RN [6]	SEQUENCE FROM N.A.	
b	171 KTFEDAKSISQHYSKRPVTPYVMPYFPDFKWINCPAQVTFDSDPAPKOTSGAALEM	230	RC	STRAIN=C57BL/6J; TISSUE=Whole body;	
b	241 MSQAMIRGMMDREGNQFTAYFLPVEETIKKRKDQEEMDYAPDDVYDVKAREYNWVK	300	RA	Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,	
b	231 MSQAMIRGMMDREGNQFTAYFLPVEETIKKRKDQEEMDYAPDDVYDVKAREYNWVK	290			

RA	Arakawa T., Bono H., Ciancini P., Fukuda S., Fukunishi Y., Furuno M., Hanagata T., Hara A., Hayatsu K., Hirao K., Hirakata T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Koijima Y., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishii K., Nomura K., Numazaki R., Ohno M., Ochiai Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shigawara A., Shiraki T., Tejima Y., Toyka T., Yamagami M., Yamamura T., Takanashi F., Tanaka T., Muramatsu M., Hayashizaki Y., Yasunishi A., Yoshida K., Yoshino M.		
RL	Submitted (IJU-L-2000) to the EMBL/GenBank/DBJ databases.		
DR	MGI: MGI:1923988; 5730511K23Rik.		
DR	InterPro; IPR017133; PF01.		
DR	PFam; PF03905; Pfam_1.		
KW	Hypothetical protein.		
FT	NON_TIR 377 MW: 43836 MW;	4ECE00D2D4BF5CEA CRC64;	
SQ	SEQUENCE 377 AA;		
Query Match	71 - 84 ; Score 1984 ; DB 2 ; Length 377 ;		
Best Local Similarity	100.0% ; Pred. No. 2.3e-86 ;		
Matches	377 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0		
Qy	1 MAPTIQTOAQREDEHRPNSHRTLPERSGVCRVKYCNSLSDIPDFDPKITYPDQNRPVQ		
Db	1 MAPTIQTOAQREDEHRPNSHRTLPERSGVCRVKYCNSLSDIPDFDPKITYPDQNRPVQ		
Qy	61 YKATSLKQHKHDLLTEPDLGITIDLINPDTYRIDPVNLIDPADEKLLEEEIQAPTSSKR		
Db	61 YKATSLKQHKHDLLTEPDLGITIDLINPDTYRIDPVNLIDPADEKLLEEEIQAPTSSKR		
Qy	121 SOOHAKVYPWRKICPEYISTEFNYGINSNEKEPKEVKIGSVRKQQTTEEEIYKDRSQTIAIE		
Db	121 SQOHAKVYPWRKICPEYISTEFNYGINSNEKEPKEVKIGSVRKQQTTEEEIYKDRSQTIAIE		
Qy	181 KTFEDAOKSIISQHNSKPKRVTPVEMPVMPVFPDFKRWINPCQAQIVFDSDPAPKDTSGAALEM		
Db	181 KTFEDAOKSIISQHNSKPKRVTPVEMPVMPVFPDFKRWINPCQAQIVFDSDPAPKDTSGAALEM		
Qy	241 M\$QAMIRGMMDDEBENQFAYFLPYETLKRRKDQEEMDYAPDDVYDKIAREYNNWVK		
Db	241 M\$QAMIRGMMDDEBENQFAYFLPYETLKRRKDQEEMDYAPDDVYDKIAREYNNWVK		
Qy	301 NKASKGYEENYFFTREGDGYYNELETRVLISKRKAGVQGSTNALLVVKERDNEKE		
Db	301 NKASKGYEENYFFTREGDGYYNELETRVLISKRKAGVQGSTNALLVVKERDNEKE		
Qy	361 LEAQEARKAQLENHEPE 377		
Db	361 LEAQEARKAQLENHEPE 377		
RESULT 9			
Q68F51	PRELIMINARY;	PRT;	407 AA.
ID	Q68F51		
AC	Q68F51;		
DT	25-OCT-2004 (TREMBLrel)	28;	Created
DT	25-OCT-2004 (TREMBLrel)	28;	Last sequence update
DT	25-OCT-2004 (TREMBLrel)	28;	Last annotation update
DE	LOC446278 Protein (Fragment).		
GN	Name=LOC446278;		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopidae;		
OC	Xenopoda; Xenopus.		
OX	NCBI_TaxID=355;		
RN	[1] SEQUENCE FROM N.A.		
RP	TISSUE-Embryo;		
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;		
RC	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.; "Genetic and genomic tools for Xenopus research: The NIH Xenopus		

Qy	454	FGSDADSEDDADSDDEDRCQAOCGSNDSDSGS-NGGQRSRSHRSASPPGSEHSAQ	512
Db	476	SRSTSRSRSRSKSSRS	535
Qy	513	ED	514
Db	536	SD	537
		RESULT 1.1	
Qy	07PKX3	PRELIMINARY;	
Db	Q7PKX3	PRT;	468 AA.
Qy	Q7PKX3;	Created)	
Db	01-MAR-2004	(TREMBLrel.	26; Created)
Qy	DT	01-MAR-2004	(TREMBLrel.
Db	DT	01-MAR-2004	26; Last sequence update)
Qy	DT	01-MAR-2004	(TREMBLrel.
Db	DB	01-MAR-2004	26; Last annotation update)
Qy	ASGP12.85	(Fragment).	
Db	GN	Name=a9GC49165 ; ORFNAMES=ENSANG00000017336 ;	
Qy	OS	Anopheles gambiae str. PEST.	
Db	CC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anophelidae; TaxID=180454 ;	
Qy	NN	L1RN	
Db	RRP	SEQUENCE FROM N.A.	
Qy	STRAIN=EST;		
Db	RCA	Anopheles Genome Sequencing Consortium;	
Qy	RRR	Submitted (MR-2000) to the EMBL/GenBank/DDBJ databases.	
Db	CCC	-1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.	
Qy	CC	EMBL: AABAB01008987 ; EAD01198.1; - .	
Db	DDR	InterPro; IPRE007133; Pfam; PF03985; Pfaf1; 1.	
Qy	DR	NON_TER	1
Db	FFT	NON_TER	1
Qy	SQ	SEQUENCE	468 AA;
Db	SQL	532	98 MW;
Qy	Query	Match	40.8%; Score 1129; DB 2;
Db	Best Local Similarity	66.8%; Pred. No. 7.4e-46;	Length 468;
Qy	Matches	223; Conservative	45; Mismatches 87; Indels 6; Gaps 5;
Qy	1	MAPPITQTAQREDCHRNPSHNRHTLPERSGVCRVKYCNSLDPDPLDKFTYPPDQNRFVQ	60
Db	5	MAPPVONGANGAD-KRP--VRQQRERSELISRVKCNTPDPLDKFTYPPFENDRFIQ	61
Qy	61	YRATSLXOKKHPLLTEDLGTTIDLNPDTYRIDPNVLLPADKEELBPIQATSSKR	120
Db	62	YNPSSLNRYTENLTHDLGTTIDLNRDLYQDISAQDPAEKLLEDEIHTPDMSR	121
Qy	121	SOQHAKVYPMRMKTYISTEFNRGISN-EKPEVKIGSVKQFETEEIYKDRDSQTAI	179
Db	122	SSRAKSVSLRKSYISTEQTRNPQTMKVEAKVGFNVKSLREBTLYMDREAQIKAI	181
Qy	180	ECTPFDQSKISQYHYSKRPVTPVEVMVFPDPFKMWNPCAQVTFPSDPAKDTGAALE	239
Db	182	ECTKEDNTKPKITHYSKPQGVTVPVIMPVFPDFAWKYPCKAQVTFPSDPAKGTNPAQIE	241
Qy	240	MMSQAMIRGMMDDEGNOQVAYFLPVEETLKKRDKQEBEMDYAPPDVYKJIAEYNNWW	299
Db	242	EMSQAMIRGMVMDSEQVAYFLPDDTLERRRDLWNETLYDEEYEYKMARVEYNNWW	301
Qy	300	KNKASKGKYBENYTFIFREGDGYYNELETYRVLKSRKAKGYSQGTNALLVVKHRDMNEX	359
Db	302	360 E 360 KSKASKGKYBENYTFVLRP-DGTYNNELETYRVLKSRQK-NAQQSNTRKLVVKHRPLNAS	359
Qy	360	E 360	
Db			

RESULT 12
Q8R91
RESULT 13
Q8R91

474 WDVQPNAND -- PGTYLVSDNGTASYL - PIPMRNLNLRKKAREGRSSDELEHFPVPSRVT 530
 QY 345 -- -TNALLVVERD - MNEKELEQEARKAOLEN - - - - -HEPEEEBEMETEKE 390
 Db 531 VRRSTIVSIEKIDSGYSSRGASSSSKMMRLUEDEGGLGRSWKHFEPQD - - - - 579
 QY 391 AGGSDEEQEKGSSSEKEGSDEHS 414
 Db 580 - - - - -ANQISDGNDYD 592

RESULT 15

Q6ZD92	PRELIMINARY	PRT;	451 AA.
ID			
AC			
Q6ZD92;			
DT 05-JUL-2004 (TRIMBLref)	27; Created		
DT 05-JUL-2004 (TRIMBLref)	27; Last sequence update		
DT 05-JUL-2004 (TRIMBLref)	27; Last annotation update		
DB Proline-rich protein-like.			
GN Name=0043BH08 .25;			
OS Oryza sativa (japonica cultivar-group).			
OC Spermatophytidae; Streptophytina; Embryophytina; Tracheophytina;			
OC Spermatophytidae; Magnoliophytidae; Liliopsida; Poales; Poaceae;			
OC Ehrhartiidae; Oryzeae.			
OX NCBI_TaxID:39947;			
[1]			
RP SEQUENCE FROM N.A.			
RA Sasaki T., Matsumoto T., Yamamoto K.;			
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.			
DR AP004460; BAC9505.1; -.			
DR InterPro; IPR007133; Pfaf1.			
DR Pfam; PF03985; Pfaf1; 1.			
SQ SEQUENCE: 451 AA; 52159 MW; 297158E4B56642AB CRC64;			
13 0% Score 358.5; DB 2; Length 451;			
Best Local Similarity 26.9%; Pred. No. 1.8e-09;			
Matches 115; Conservative 80; Mismatches 180; Indels 53; Gaps 17;			
Qy 17 PNSHR--TIPERSGVVCRVKYCNSLDPIDPDKFTYPFDONRFVQYKATSLEKOKHD 73			
Db 47 PNAERFENLKECPPTFLCKHKERNELDPSSQLKWLPLNOKDRTTKYRITSLEKNYIPK 106			
Qy 74 LITEPDLGVTIDLINPDTYRIDP-NVLLDADEKLL_EEIQOAPTS---KRSQQHAKV 127			
Db 107 MIVPEDLGPLDLDMSVNTPVQPPMAPPEDBELLRDDEVLTPVKDGIRKKERPTDKG 166			
Qy 128 VPMRKRTEYI---STEFNRVGI---SNEKPEVKIGSVKQFTEBEIYKDRDSQITAE 180			
Db 167 MSWLVKQYISPLSTDAAKMSITEQAKERREPSREG --RNTFLEN - INDREKQIAKE 221			
Qy 181 KTFEDAOKSIQHYSKPRVTPVVMPVFDOWINPCAQVIFDSDPAPKDTSGAALE- 239			
Db 222 DSFR-AAKSRRPHOTKRCMEAWWLPLPFDYRDDOFVMNFDGDET-AUSEQTNKLER 279			
Qy 240 -----MSQAMTRGMMDDEGNOFVAYFLPVESTLKERKRDQEEMDYAPDDVVDY 289			
Db 280 SERDECBSRAVKSFLYNGSDPAKQERFLATVNPSPHLSQDDETEDIOYS----- 332			
Qy 290 KIAREYNNWNQKNSKQGYENTFFIFREGGTYNELETYRVLKERAQAVQSGTNALL 349			
Db 333 -WIRETHMEVRGD DDDDPPTYLVTF-DDDGAKYLPLPTKLVIQKCKAKEG-RSGDE -- 385			
Qy 350 VVGHDRDNNEKELEQEARKAOLENHEPEEEEBEMETEKEAGGSDEOEKGSSSEKEGS 409			
Db 386 -TEHFVPSRITYSRTAHHGMMEHGESSMH ---ENLKQRSSVDDLYDHPKHSRVED 440			
Qy 410 EDEHSGSE 417			
Db 441 MDQSGDE 448			